DOCUMENT RESUME

ED 061 430 VT 014 837

TITLE	Evaluation--An Objective Approach.
INSTITUTION	National League for Nursing, New York, N.Y.
REPORT NO	NLN-Pub-16-1446
PUB DATE	72
NOTE	53p.; Report of the 1971 Workshops of the Council of Diploma Programs
AVAILABLE FROM	National League for Nursing, 10 Columbus Circle, New York, N.Y. 10019 (Pub. No. 16-1446, $2.00)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	Conference Reports; Course Evaluation; *Course Objectives; Curriculum Development; *Curriculum Evaluation; Evaluation Methods; Evaluation Techniques; *Health Occupations Education; *Nursing; Student Evaluation; *Workshops

ABSTRACT

A total of 688 participants representing 356 agencies attended one of five 2-day workshops, which constituted the third phase of a 3-year curriculum development program and were designed to: (1) recognize the constructive aspects of evaluation in the educational process, (2) establish positive perceptions of evaluation for subsequent use in the assessment of students, (3) determine the relationship of evaluation tools to the objectives of specific nursing courses, and (4) resolve to use course objectives in the development of evaluation tools and techniques. Utilizing the 1969 and 1970 reports as a base, workshop objectives were met through group work and reactions to major presentations by H.H. Flitter, L. Litwack, L.M. DeRidder, and M.E. Katzell who spoke on "Evaluation: An Objective Approach." Texts of the major presentations are included, and workshop materials are appended. The 1969 and 1970 reports are available as ED 039 356 and ED 050 274, respectively. (SB)

ED 061 430

EVALUATION— AN OBJECTIVE APPROACH

Report of the 1971
Workshops of
The Council of
Diploma Programs

NATIONAL LEAGUE
FOR NURSING • 1972

VT014837

ED 061 430

EVALUATION--AN OBJECTIVE APPROACH

Report of the 1971 Workshops of
the Council of Diploma Programs

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

NATIONAL LEAGUE FOR NURSING
Department of Diploma Programs
10 Columbus Circle, New York, New York 10019

1972

Publication Number 16-1446 Price: $2.00

Copyright, 1972, by

NATIONAL LEAGUE FOR NURSING

All rights reserved. This book, or parts thereof, must not be reproduced in any form without permission of the publisher.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
National League for Nursing
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

PRINTED IN THE UNITED STATES OF AMERICA

CONTENTS

INTRODUCTION 1

EVALUATION--AN OBJECTIVE APPROACH

Presentation by:

Hessel H. Flitter 8

Lawrence Litwack 15

Lawrence M. DeRidder 26

Mildred E. Katzell 38

APPENDIXES

A. Master Agenda 47

B. Work Sheet for Course Objectives 48

C. Work Sheet for Clinical Laboratory Evaluation Tool 49

INTRODUCTION

The 1971 workshops, sponsored by the Council of Diploma Programs, constituted the third phase of a three-year program planned by the Program Committee of the Council. These workshops, which had as their theme Evaluation--An Objective Approach, were, in effect, an outgrowth of the activities of the 1969 and the 1970 workshops, which dealt with the subjects Level Objectives--Development and Use in the Curriculum and Keeping on Course, respectively. The Committee therefore suggests that this report be studied in conjunction with the two previous workshop reports.

Prior to the 1971 workshops, a master agenda (see Appendix A) was drawn up and mailed, together with the workshop objectives, to each agency member of the Council. The objectives, established by the Committee, were the following four:

1. Recognize the constructive aspects of evaluation in the educational process.
2. Establish positive perceptions of evaluation for subsequent use in the assessment of students.
3. Determine the relationship of evaluation tools to the objectives of specific nursing courses.
4. Resolve to use course objectives in the development of evaluation tools and technics.

In view of the nature of the objectives, the Committee recommended that these guidelines be considered when selecting representatives:

1. Representatives should be instructors who are involved in planning and implementing all aspects of a clinical course in nursing.
2. If feasible, at least one of the representatives should have attended either a 1969 or a 1970 workshop, or both.

The Committee also requested that each representative, when filling out the registration card, indicate on the list provided the subject area to which the course she was teaching was most closely related. The subject areas listed were fundamentals of nursing, medical-surgical nursing (the initial course, the midway course, and the advanced course), maternity nursing, nursing of children, maternal and child nursing, and psychiatric nursing.

If--before launching into a description of the 1971 workshops--we take a quick look at the workshops of the year before, we see that they were so structured that they culminated in a presentation of evaluation in general--the importance of evaluation and the methods of collecting data for evaluating curricula, courses, and students. Reviewing the 1970 workshops and the subsequent comments and suggestions of the participants, the Program Committee came to the conclusion that what the members really wanted was to concentrate on a single aspect of evaluation, an aspect that was of major concern to them--the evaluation of the student in the clinical laboratory. In order to make sure that all the points vital to such an evaluation would be covered at the workshops, the Committee requested in advance that each of the speakers invited to present an introductory paper on the subject include the following areas:

1. The beliefs influencing evaluation.

2. The application of the basic concepts of evaluation to course evaluation and the attainment of course objectives by the student.

3. Ways of using course objectives to evaluate the student in both the classroom and the clinical-laboratory aspects of a course.

4. Some current trends and practices in determining a grading system.

The Program Committee's plans were duly put into operation at each of the five workshops. (For a list of the cities in which the workshops were held, see page 6.) At each of them, one paper, focusing on the evaluation of students in the laboratory area, was read, and a general discussion followed. In all, four papers were presented (one of the papers was read at two workshops); all are included in this publication.

When the time for the work sessions arrived, the participants broke up into the groups to which they had been assigned according to the subject area they had specified on their registration cards. Each representative had been requested to bring to the meeting:

1. A list of the objectives of the particular course she was teaching.

2. The evaluation form or tool used to evaluate the students' clinical performance at the end of the course, including, if any, written instructions or guidelines for its use.

3. The evaluation form or tool completed for two students at the end of the course.

These materials, together with a set of directives prepared by the Committee and distributed at the workshop, formed the basis upon which each group set about its task. We are reproducing the directives below:

> You are a member of the faculty of WENS School of Nursing, and you have been appointed to serve on an ad hoc committee of the curriculum committee charged with the task of determining the relationship of evaluation tools to the objectives of a specific course in nursing.
>
> The curriculum committee, in its charge to the members of the ad hoc committee, directed their attention to the determinations on which the planning and evaluation of the curriculum were based. They are:
>
> 1. WENS school objectives identify behaviors of the new graduate that result from participation in the school's services and curriculum offerings.
>
> 2. The curriculum, or last-level, objectives of WENS identify the behaviors of the graduate that result specifically from participation in the total program of study (curriculum).
>
> 3. The WENS faculty, following a recommendation from the curriculum committee, established levels for the curriculum and adopted objectives for these levels.
>
> 4. The level objectives were used to plan the curriculum by placing all courses in appropriate levels, identifying course objectives, evaluating the courses placed within each level, and evaluating the total curriculum.

5. The objectives for each course were developed to show a relationship to (a) level, curriculum, and school objectives and (b) course, unit, and lesson objectives. The diagram below was found by the curriculum committee to show the relationships.

School objectives ⟷ curriculum objectives ⟷

level objectives ⟷ course objectives ⟷ unit

objectives ⟷ lesson objectives.

In order to expedite the task assigned to the committee, you and the other members agreed to use as references certain definitions of terms approved by the faculty organization and some guidelines forwarded from the curriculum committee.

Definitions*

School: An entity that is controlled by an identifiable group; has students enrolled for a specific purpose; has instructional personnel; has other appropriate personnel; provides services to the student body, faculty, and other personnel; has educational facilities and resources; and offers a planned curriculum.

Curriculum: The course offerings that make up the program of study.

Course: The organized subject matter and related activities, including theoretical and laboratory experiences planned to achieve specific objectives and offered within a specified time period.

Course unit: A major subdivision of a course.

Objective: A statement describing a measurable proposed change in the behavior of the student upon completion of a learning experience. (The specificity of objectives and, in some instances, the number of related objectives increase in progression from school to curriculum to level to course to unit to lesson.)

Guidelines*

Course objectives should clearly differentiate one course from another course.

Course objectives should be used to plan, implement, and evaluate all aspects of a particular course.

Course objectives should be subdivided into unit objectives that are in turn subdivided into lesson objectives.

Objectives should be used to identify learning experiences that will bring about the achievement of the desired behavior expected of the student.

A sufficient variety of experiences representing the desired behaviors should be planned for the student by the instructor.

*For use only at workshop.

Instructors should assume responsibility for planning and participating in all aspects of the course.

Objectives of the course, unit, and daily lessons should be shared with the students.

Students should participate in a variety of ways and activities that relate to planning, implementing, and evaluating the course.

Students should be involved throughout the course in activities related to self-evaluation.

Instructors should plan for periodic evaluation of the student at given intervals throughout the course.

Evaluation of the student's clinical laboratory performance should be but one aspect of evaluation when determining the student's attainment of course objectives.

Certain evaluation activities and methods are unique to the clinical laboratory experiences that have been planned to assist the student to attain course objectives.

The instructor should engage in a variety of activities in order to collect data on a day-to-day basis regarding the student's performance in the clinical laboratory. Some of the activities are:

1. Observing the student's planning and carrying out of the nursing care of a patient.
2. Discussing with the student the achievement of the lesson and unit objectives.
3. Evaluating the student's nursing care plans, nursing notes made on the patient's chart, and other written assignments in the light of the objectives.
4. Evaluating the depth of the student's contributions to discussions and planned conferences.
5. Collecting from patients, families, other nursing personnel, other health workers, et cetera, pertinent comments regarding the student's performance.
6. Keeping anecdotal records related to the above activities.

The purpose of the group work session was to have the course objectives and the evaluation tools submitted by each individual considered by the group. First, the participants exchanged their lists of objectives with one another so that everyone's list should come under the scrutiny of another person, thus facilitating an objective approach. A work sheet was used as a guide to the evaluation (see Appendix B). When all the individual evaluations had been completed, each person made her observations known to the others, and a general discussion ensued.

The next step was to review the evaluation tools (a blank form and a completed form) in relation to the course objectives. Each tool was examined, again with the aid of a work sheet (see Appendix C), by the same person who had gone over the course objectives to which it related and, once more, individual observations were shared and discussed.

Due to the time limit, however--an hour for each evaluation--only selected objectives

on each participant's list could be considered. The process of evaluating course objectives and their use in the evaluation tool were the major focus of the discussion.

At the close of the group session, the members were asked to select examples of information given under each column of the work sheets that they thought should be discussed with the entire assembly. Where the same subject areas were selected by more than one group, the reports were incorporated into one and presented by one person.

The group reports at each of the five workshops revealed the following commonalities:

1. Faculties are making significant progress in stating course objectives in behavioral terms.
2. Fewer objectives fail to identify the course content. For example, such nonspecific statements as "utilizes available community resources in planning care of a patient" appear less frequently.
3. Some objectives still fail to identify clearly measurable behaviors--for example, "grows in the ability to use self therapeutically."
4. Some stated objectives do not seem to be pertinent to the course content and the course experiences--for example, "assumes responsibility for professional and personal development."
5. The objectives and the titles of some courses appear inconsistent--for example, advanced medical-surgical nursing often does not include content related to care of medical-surgical patients but focuses on team nursing, trends in nursing, et cetera.

A wide variation in the tools used for evaluating students in the clinical area was evident. Some schools use forms prepared by an outside agency; others have developed the same tool for use in all clinical nursing courses. Then again, there are schools that use level objectives, with specific behaviors identified, and schools that use course objectives. Still others prefer a combination of approaches. The rating method also varies from tool to tool, as do provisions for narrative descriptions. Many groups reported that the tools their schools were using were currently under revision; from what they told, the fact emerged that there was a definite trend toward making course objectives the basis of evaluating students. It was also pointed out that not enough guides had been devised for the use of tools.

The consensus of the group reports was that each faculty must determine for itself (1) whether or not the objectives of a particular course meet the criteria established by the faculty and (2) how the objectives can be effectively used in the development of an evaluation tool.

The final session of each workshop included three case presentations illustrating how three different faculties shifted to the use of course objectives in the evaluation of clinical performance. The accounts stressed the manner in which the faculties effected the shift and the factors that facilitated or hindered their progress.

Each presentation exemplified a different approach, different circumstances, and different outcomes, yet it was evident from the reactions of those assembled that most of the points made struck a familiar chord. The Council is planning to publish the presentations--a total of 15--sometime in 1972.

Before leave-taking, the participants were urged to initiate the evaluation procedure

elaborated at the workshop in their home schools in cooperation with their own faculties. Also, it was suggested to those participants working in areas where more than one school of nursing was located that they organize a workshop similar to the one they had just attended, inviting representatives of the local faculties to take part.

The reports, observations, and discussions at all the workshops reflected the substantial progress the faculties of diploma schools of nursing are making in the use of objectives--in curriculum and course development as well as along various avenues of evaluation.

The table below lists the cities in which the five workshops were held and shows the number of agency members represented and the number of registrants at each:

Workshop	Agencies	Registrants	
		Representatives	Guests
Fort Wayne, Indiana	84	162	3
Denver, Colorado	54	100	3
Philadelphia, Pennsylvania	94	181	2
Hartford, Connecticut	89	183	2
Jacksonville, Florida	35	62	1
Total	356	688	11

The 356 agency members represented at the workshops constitute 67 percent of the total membership of the Council. The 11 guests were representatives of state boards of nursing.

The dates of the workshops and the names and positions of the Program Committee members who presided, the principal speakers, and those who delivered the case presentations follow:

FORT WAYNE, INDIANA, September 23-24.

Presiding: E. Wanda Quay, Director, School of Nursing, Riverside Hospital, Toledo, Ohio.

Principal Speaker: Lawrence Litwack, Ed.D., Chairman, Counseling and Personnel Services Education, Kent State University, Kent, Ohio.

Case Presentations by: Annajean Boukamp, Instructor, Medical-Surgical Nursing, Hackley Hospital School of Nursing, Muskegon, Michigan; Erliene Clayton, Instructor, Child Health Nursing, St. Elizabeth School of Nursing, Lafayette, Indiana; Alice M. Jordan, Instructor, Maternity Nursing, Community Hospital of Springfield and Clark County School of Nursing, Springfield, Ohio.

DENVER, COLORADO, October 4-5.

Presiding: Dorothy D. Rademacher, Instructor, Emanuel Hospital School of Nursing, Portland, Oregon.

Principal Speaker: Hessel Flitter, Ed.D., Associate Dean for Nursing, Chico State College, Chico, California.

Case Presentations by: Tommie Crooks, Instructor, Fundamentals of Nursing, Charity Hospital School of Nursing, New Orleans, Louisiana; Todette Holt, Instructor, Medical-Surgical Nursing, St. Luke's Hospital School of Nursing,

Fargo, North Dakota; Katharina Wood, Instructor, Medical-Surgical Nursing, The Finley Hospital School of Nursing, Dubuque, Iowa.

PHILADELPHIA, PENNSYLVANIA, October 14-15.

Presiding: Marjorie J. Montanari, Administrative Assistant in Charge of Patient Services, Citizens General Hospital, New Kensington, Pennsylvania.

Principal Speaker: Mildred E. Katzell, Ph.D., Director, Division of Measurement and Evaluation Services, National League for Nursing, New York, New York.

Case Presentations by: Viola Diehl, Coordinator, Medical-Surgical Nursing, The Allentown Hospital School of Nursing, Allentown, Pennsylvania; Sister Louise Grundish, Assistant Director, Pittsburgh Hospital School of Nursing, Pittsburgh, Pennsylvania; Bonnie Owens, Instructor, Fundamentals of Nursing, Washington Hospital Center School of Nursing, Washington, D.C.

HARTFORD, CONNECTICUT, October 28-29.

Presiding: Natalie L. Petzold, Director, School of Nursing, Massachusetts General Hospital, Boston, Massachusetts.

Principal Speaker: Mildred E. Katzell, Ph.D. (See Philadelphia Workshop.)

Case Presentations by: Jean Bieszad, Instructor, Maternal and Newborn Nursing, St. Francis Hospital School of Nursing, Hartford, Connecticut; M. Evelyn Guertin, Medical-Surgical Nursing I Coordinator, Meriden-Wallingford Hospital School of Nursing, Meriden, Connecticut; Louise Phall, Educational Chairman, St. Vincent's Hospital School of Nursing, New York, New York.

JACKSONVILLE, FLORIDA, November 4-5.

Presiding: Mary Catherine King, Director, School of Nursing, St. Vincent's Hospital, Birmingham, Alabama.

Principal Speaker: Lawrence M. DeRidder, Ph.D., Professor and Head, Department of Educational Psychology and Guidance, The University of Tennessee, Knoxville, Tennessee.

Case Presentations by: Louise B. Duncan, Assistant Director, Piedmont Hospital School of Nursing, Atlanta, Georgia; Margaret Heins, Director, School of Nursing, St. Mary's Memorial Hospital, Knoxville, Tennessee; Elizabeth Parris, Instructor, Medical-Surgical Nursing, Greenville General Hospital School of Nursing, Greenville, South Carolina.

EVALUATION--AN OBJECTIVE APPROACH

Presentation by Hessel H. Flitter

One aspect of the revolution that is said to be taking place in our educational system today is that of criticisms by students of the "Mickey Mouse" character of the system. By "Mickey Mouse," they seem to mean that the educational process consists not of preparation for a life-role but of a sort of "Tweedledum-and-Tweedledee" battle over pedantic trivia that are wholly unrelated to their life-role. Much of the evidence that the student generation uses to support its criticism is based upon what it sees as the meaninglessness of tests, grades, and "merit badges" that are supposed to predict performance in life-roles. In short, the students accuse the educational system on the basis of the shortcomings of its evaluation procedures.

A concomitant feature of the "student power" movement is that students no longer attribute authority to the teacher just on the basis of role. Students are quick to question the value of the educational experiences provided them and especially of evaluative devices and methods. It is not unusual today for a director of a nursing education program to be faced, in person or in a letter, with a student who contests the competencies of one of her teachers. Almost invariably, a basis for the criticism is, "She doesn't grade me fairly." Moreover, when the "worm turns," there is apt to be a reaction on the part of the teacher that not infrequently makes her more liable to the student's criticisms. Let me explain.

Compared to those in other disciplines, we are rather fortunate. We have a much clearer conception of the life-role of the nurse than the general educator has of the life-role of the good citizen. And this we do know: nursing is a complex behavior involving variables that include skill, knowledge, attitudes. Nursing is a multifarious, many-faceted phenomenon. Certain of the variables can be measured accurately, but for others we have at best only crude measuring devices. For instance, the ability to recall the names and shapes of instruments to be placed on a tray for a particular procedure can be measured with great accuracy. The ability to put a patient in a frame of mind that facilitates the procedure is measured crudely at best, and the very basis of the evaluation is apt to be questioned.

Now to return to the reaction of the teacher whose evaluation has been questioned by the student. If the teacher is frightened and insecure, she is apt to limit her evaluation to those items that have a high degree of accuracy and that include perhaps a single measurement. Very often, her reaction is self-defeating. Because the life-role of the nurse is multifarious, the student is even more justified in criticizing the evaluation as being unrelated to the objectives.

The answer to this dilemma seems to be that every teacher must think through rather deeply what he believes about the nature of evaluation. He must arrive at a set of beliefs about the purpose and the nature of the evaluative process of which he is convinced and to which he is committed. His security in situations such as I have described comes

---

Hessel H. Flitter, Ed.D., is Associate Dean for Nursing, Chico State College, Chico, California.

from his philosophy of evaluation. My years of experience as a teacher and administrator lead me to the conclusion that there is no single philosophy of evaluation that is a ready-made fit for every teacher. To spring a surprise quiz on the students would be thoroughly inconsistent with the philosophy of evaluation of many teachers. Yet, I can conceive of a teacher whose surprise quiz would be accepted as a constructive educational experience.

I invite you for a moment to consider these two opposing points of view, or philosophies, of evaluation. Which one has greater appeal for you? With which do you agree in whole or in part? My goal is not that of setting up two armed camps of opposing philosophies of evaluation. I strongly suspect that most of you will not be most comfortable at either pole. So as to have a name that we can use to identify the poles, let us call one pole the unitary approach and the other pole the holistic approach.

Those who support the unitary approach believe that any meaningful evaluation consists wholly of measurements that are universally acceptable. They believe that we can, if we work hard enough, break down even complex behavior into its constituent operations which can then be measured with complete objectivity. If we want to measure a quality such as the "ability to put a patient at ease," we can separate putting a patient at ease into its component parts and devise situations that would measure the ability to do so--and all examiners, given the same examinee, would obtain the same score. But what about those qualities, skills, and abilities for which we do not have such accurate measurement? Those at the unitary pole would merely say, "If there are no proven evaluative devices, don't use unproven ones." But the refusal to attempt to use an unproven tool to measure a quality doesn't mean that one ignores the quality. Rather, one can use proven devices that have been shown to measure something that is related to the quality in question. For example, suppose that we do not yet have a "putting-a-patient-at-ease" test but that it has been proven scientifically that nurses who can recall certain concepts from psychology do in fact put the patient at ease. Then we would use the related proven evaluative device.

The beauty of the unitary approach is that by demanding a high degree of accuracy and reliability we are able to use the evaluations that we do make to compensate for some of the current gaps in evaluative devices. Those at the unitary pole visualize the evaluator as a highly skilled person, capable of complete objectivity, whose evaluations are in complete agreement with other highly skilled evaluators.

Now let us swing to the holistic pole. Those who cluster here see the evidence for evaluation as having no limits. They, also, look for operational objectives, but they would not discard evidence pertaining to any of these objectives simply because the evidence is crude or unproven. They would give preference to the more accurate of two devices that evaluate the same quality, but they would not be unduly influenced by a test result that was "carried out to three decimal places." In fact, they tend to be suspicious of the numerical artifacts of measurement in the case of human behavior. They are greatly impressed with the many variables that operate in human behavior and would have no qualms in giving quantity of variables preference over the exact and exacting measure of a single variable.

Those who cluster at the holistic pole are totally flexible regarding the nature of both evaluations and evaluators. They believe that at least a part of the evaluations should be made in the real-life situation by the nonteacher. They see the student as having superior competency to evaluate her own progress and would not hesitate to include an evaluation by the patient or the maid if it pertains to the operational objectives being considered.

The holistic evaluators see an evaluation that is communicated as a "3.25" average as working against the educative process in that it predisposes the student to equate evaluation with unrealistic educational shibboleths rather than with stimulation of self-development and improvement.

Now to get down to the "nitty-gritty" of the workshop--namely, how can evaluation assist you in reaching the objectives of your class, your course, or your program? Although previous presentations at the workshops during the past two years gave excellent coverage of objectives--what they are and how they are developed--the inseparable nature of objectives and evaluation warrants a little reflection on some of the ideas that were previously discussed. One of the themes of this series of workshops is that evaluation is an integral part of the teaching-learning process. If we don't integrate evaluation, our entire educational process is apt to disintegrate. If the cart (evaluation) doesn't follow the horse (objectives), we wind up going nowhere because our horse has become a cart. Let me explain what I mean.

We all state blithely that we are evaluating the objectives of our courses, but from the students' point of view, the course evaluation is a much more real statement of what we expect of them than is a pretty statement of objectives that appears in the course outline. Actually, evaluation is a part of the educational process, whether or not we intend it to be so. If, for example, we stress outside readings in our objectives but test only on materials covered in lecture, the student quickly learns (1) what our real expectations are and (2) how to use his study time so that it profits him most on the day of reckoning. From the student's point of view--or if you will, from the point of view of reality--it is meaningless to say objectives should determine evaluations. Rather, the evaluations show the true nature of the objective.

Objectives tell us not only what should be taught and learned but also what should be evaluated. Your objectives determine your evaluation every bit as much as your shopping list determines what you buy in the supermarket. Therefore, I strongly suggest that before you go to your educational supermarket, you consult a sort of superchecklist of all the types of items you might be looking for. Just as your shopping list will help you in determining which department will provide you with a needed item, so a superchecklist will help you in determining the type of evaluative device you need. You will recall from previous workshops that a handy superlist is a taxonomy of objectives (a misleading name, I think, because a taxonomy makes the teacher's task less taxing).

Let us consider one somewhat simplified taxonomy of objectives and how it relates to the basic concepts of evaluation.

A. Knowledge:

- A.1 Terminology
- A.2 Specific facts
- A.3 Generalizations
- A.4 Concepts
- A.5 Methods and procedures

B. Understandings--Ability to:

- B.1 apply knowledge in new situations
- B.2 interpret cause and effect
- B.3 explain methods and procedures

C. Thinking Skills--Ability to:

C.1 generalize from given data
C.2 recognize assumptions underlying data
C.3 recognize limitations of data

D. General Skills:

D.1 Laboratory skills
D.2 Communication skills
D.3 Computational skills
D.4 Study skills
D.5 Social skills

E. Attitudes:

E.1 Social attitudes
E.2 Scientific attitudes

F. Interests:

F.1 Personal interests
F.2 Professional interests

G. Adjustments:

G.1 Social adjustments
G.2 Emotional adjustments

We will return to our shopping list shortly, but first, the prudent consumer needs additional knowledge if he is to invest wisely. He needs basic concepts about the nature and quality of his purchase. Let me share with you some basic concepts that I believe should go into the selection of each evaluative experience.

1. The evaluation should be a valid measurement of the objective. It should be on target, so to speak. If the objective is the attainment of a specific behavior, the evaluation should measure that behavior, not another behavior.

2. The type of evaluation should be appropriate for measuring the objective. If, for instance, the objective is ability to operate a suction machine, an appropriate evaluation would entail its operation.

3. The evaluation should include a significant sample of the behaviors that the objective implies. It should include a sufficient number of behaviors that are critical evidences of the attainment of the objective.

4. As far as possible, the evaluation should show the degree to which the objective has been achieved. Excepting all-or-none situations, the evaluation should indicate the degree of mastery of a particular objective.

5. The evaluation should contribute to the educative process. Contributions include stimulating and motivating students, diagnosing students' difficulties, and improvement of instruction.

Now, if we combine the shopping list of objectives and basic concepts of evaluation, we are ready to be discriminating shoppers in our evaluation supermarket. Let us sup-

pose, for example, that we are conducting a course in the nursing of children. On our shopping list (taxonomy), under A.1, we have included under "Terminology" many terms that the nurse must understand in order to communicate about the care of the child. If our objective is ability to identify a term with its definition, constructing a valid device is rather simple. The traditional paper-and-pencil test situation seems appropriate. If we ask the student to identify a number of terms, we can show the degree of mastery of the subject. One problem may be the selection of a sample of terms that are significant. Here we may use as our criterion for selection such factors as the frequency with which a given term is used in communications when caring for an ill child or the degree to which a given term is critical to the understanding of the content of the course. In our selection, we may also be guided by the possible contributions of the evaluations to the educative process. Perhaps we may want to use this test to assist in diagnosing students' difficulties in comprehending what they read about the nursing of children or we might want to evaluate the effectiveness of our strategy for vocabulary building.

If, however, we are at point G.2 on our shopping list--namely, emotional adjustment, and we have an objective related to the student's adjustment to the child with a behavioral problem, our approach to selecting an evaluative device may be totally different. We may reject any paper-and-pencil test and base our evaluations on the student's and teacher's anecdotal records of the student's reactions when she was assigned to the care of the child with deviate behavior.

So far, we have discussed evaluation in general and how your personal philosophy of evaluation affects the devices you will probably use in measuring outcomes of the teaching-learning process. We have also considered the relationship of objectives to the evaluative process and offered a few examples of the relationship. The last topic that I would like to explore with you is one that is timely in terms of the changed values students hold with respect to education. I am referring to the topic of grading. Some faculty members as well as some students would discard all grading systems and assign a passing grade to all students who attend class regularly and turn in their assignments. At the other extreme are those faculty members and students who hold the belief that a grade is the most important outcome of an educational experience, whether or not learning has taken place. In both of these views, it is assumed that all of the grading is done by the teacher. Another point of view is self-grading, or self-appraisal. Students establish their own goals, either alone or with faculty, and evaluate their own progress toward these goals. This system is reported in use in evaluating clinical laboratory experiences particularly.

Among faculty members who accept the need for a grading system in which the teacher determines the grade, there exist three predominant points of view. These are the so-called individual-standard, the fixed-standard, and the group-standard systems. The first of these, the individual-standard, is sometimes called ability grouping. It assumes that one can determine each student's relative standing with respect to ability to acquire knowledge. A standard of achievement appropriate to a particular student's potential to learn is then established. A lesser degree of achievement is expected of the less gifted than of the brighter student. This system is reported in use in evaluating clinical performance in some schools. For example, a student poor in planning nursing care in her first assignment who later excels in this aspect of clinical performance may receive a higher clinical rating than one who begins at a higher level and does not go beyond the initial level.

The most traditional and most widely used system is that labelled <u>fixed-standard</u>.

This sytem utilizes percentage grades based on the actual number of items successfully passed in relation to the total number of items on the test. Although the term fixed-standard is used, in actual practice the only thing fixed is the meaning of percentage. The standard can easily become variable if the teacher varies the difficulty levels of tests. An easy test yields higher percentage grades and a difficult test, lower percentage grades, other things being equal.

A popular system today is the group-standard. In this system a student is compared to others in the group and not to a fixed competency level. Grades are assigned by first ranking the students by their raw scores and then deciding on some way of assigning letter grades to the raw scores. The group standard allows for flexibility of standards to fit the intellectual level of each group. The question is sometimes asked, "Should I use a normal curve in assigning grades to my class?" One does not have to be an expert statistician to know that the normal curve is not applicable to a small class whose members have been selected on the basis of ability. A class whose makeup is highly homogeneous in terms of scholastic ability may not merit any failures. When a normal curve is used, 7 percent of the class receives a failing grade by virtue of the distribution of the curve. A very practical method used by many faculty members is that of natural breaks. For example, in a raw-score distribution ranging from 48 to 127 and involving 35 students, the range of A's was from 120 to 127; B's, from 96 to 108; C's, from 71 to 84; D's, from 61 to 67; and F's, from 48 to 53. There were 4 A's, 8 B's, 14 C's, 6 D's, and 3 F's. Had a normal distribution been used there would have been 2 A's and 2 F's, 8 B's and 8 D's, and 15 C's. The difficulty in this method arises when there are no natural breaks in the distribution of grades.

Most faculty members say that their least rewarding job is assigning a grade to a clinical nursing experience. The dissatisfaction expressed may be due to a lack of measures that are universally accepted and objectively scored. Many schools have developed rating scales--some have been arbitrarily constructed and others have been tested over many years and validated against other measures, such as peer evaluations by students. The difficulty with grading a clinical experience may lie in its multifarious character. In the August 1971 issue of Nursing Outlook, a so-called creative examination of clinical experience was described. Students observed a film and were then questioned by means of a paper-and-pencil test on various aspects of the film. The authors admit that the method does not give evidence of the students' ability to implement the patient care that they can verbalize well on the paper-and-pencil examination. However, there is an advantage to a filmed test in that it reduces the variation that is present in the test situation when the test is an actual clinical situation.

If we could break down clinical experiences into behavioral objectives as clearly as we can classroom aspects of nursing, we would have the beginning step for developing evaluative devices to appraise these experiences. But this would be only the beginning. If we subscribe to the holistic approach to evaluation, we would have qualms about ascribing numbers or even letter grades to some aspects of nursing. It is like trying to grade a stew. If you grade the carrots B, the potatoes C, the meat A, and the broth D, do you average the whole and call it a C-plus stew?

Bibliography

Bloom, B.S., ed. Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York, Longmans, 1956.

Bormuth, J.R. On the Theory of Achievement Test Items. Chicago, University of Chicago Press, 1970.

Ebel, R.L. Measuring Educational Achievement. Englewood Cliffs, N.J., Prentice-Hall, 1965.

Green, J.A. Measurement and Evaluation. New York, Dodd, 1970.

Heins, M., and others. Evaluation and Prediction of Clinical Performance in a School of Nursing. Knoxville, Tenn., St. Mary Memorial Hospital School of Nursing, 1971.

Krathwohl, D.R., and others. Taxonomy of Educational Objectives, Handbook II: Affective Domain. New York, McKay, 1964.

Mager, R. F., and Piper, P. Analyzing Performance Problems. Belmont, Calif., Fearon-Lear Siegler, 1970.

Marshall, J.C., and Hales, L.W. Classroom Test Construction. Reading, Mass., Addison-Wesley, 1971.

National League for Nursing. The Construction and Use of Teacher-Made Tests, 2d ed. New York, the League, 1965.

EVALUATION--AN OBJECTIVE APPROACH

Presentation by Lawrence Litwack

Basic Assumptions

As I thought through the material to be covered in this presentation, I found myself recalling the concerns and problems shared by numbers of nursing educators over the last eight years. I reviewed the material appearing in the professional journals and reread the summaries from last year's series of workshops. As I began to organize the wealth of material available, I found myself returning to certain basic assumptions that form the foundation for the principles of evaluation I hope to share with you. None of these assumptions are necessarily new, but they need restating in order for us to start from a common frame of reference. They may be summarized as follows:

1. Clinical evaluation is the most important dimension in the process of evaluating student learning.
2. Evaluation is a process, rather than a series of single events.
3. Mandatory knowledge and skills should be clearly identified, completely communicated to students, and adhered to as a standard.
4. Evaluation is a tool in the educational process, never an end in itself.
5. Group evaluation is always better than individual evaluation, unless the individual has unique knowledge about the student.
6. Evaluation tools too often impede rather than enhance learning.
7. External evaluation tools, such as standardized tests and state board examinations, are of less importance in the evaluation process than the professional judgment of the faculty within a school of nursing.
8. The student must share the responsibility for objective and valid evaluation of her progress.
9. Too many faculty members have lost sight of their primary responsibility, which is to educate, not to evaluate.

By no means will you agree with each of the assumptions stated above, but I hope you will keep them in mind as I talk so that we can return to them later if you wish.

Overview of Clinical Evaluation

The process of evaluation in the clinical area is perhaps the most difficult task faced by most nursing educators. Therefore, I will concentrate on this area in my presentation.

Lawrence Litwack, Ed.D., is Chairman, Counseling and Personnel Services, Department of Graduate Education, Kent State University, Kent, Ohio.

By clinical evaluation, I refer to the evaluation of all educational experiences that occur outside a formal classroom setting. Part of this difficulty lies in the fact that clinical evaluation too often is highly ambiguous in nature and is particularly vulnerable to problems of subjectivity and instructor bias. The other major difficulty rests on the premise that clinical evaluation is perhaps the most important dimension in the assessment of student progress and learning, since under no other circumstances is the student as closely supervised,or is such a detailed evaluation made of her clinical competence.

Certain problem areas serve to highlight the current situation. The first revolves around discrepancies between faculty expectations of students and the expectations of nurse practitioners. The second deals with the discrepancy between students' objectives and those of the nursing faculty. In the first area, the resulting conflicts frequently leave the student caught in the middle. On the one hand, she is evaluated by what she feels are the idealistic standards of the faculty. On the other, she constantly views around her the wide divergence in the techniques and practices used and the attitudes expressed by graduate staff nurses. The resulting obvious contradictions frequently lead to the disillusionment felt by students at some stage of their program. This same conflict was found in student-faculty attitudes and expectations. In both areas, the need for a thorough orientation of nursing staff and students prior to the commencement of any clinical experience seems to be clearly indicated. If we intend to make the clinical experience a valued and integral part of the educational process, then there needs to be initial agreement on the objectives to be reached and the roles of faculty, students, and nursing staff as contributors to the attainment of these objectives.

The Qualification Process

Before one can begin to evaluate, it is necessary to clearly designate the goals to be reached and the methods to be used in assisting students to reach these goals. Evaluation does not occur until later, when we wish to assess each student's progress towards achieving the preestablished goals. Therefore, it seems logical in any discussion of clinical evaluation to proceed in an orderly fashion from the initial setting of goals to the final evaluation conferences and grading procedures. This process is designed to enable both the students and the faculty of a school of nursing to learn and grow together in a mutually helpful relationship, with the faculty members sharing with the students their skills, knowledge, understanding, objectives, et cetera, and the students sharing with the faculty their concerns, their reactions to the learning process and the educational environment, and their self-assessment of their progress and areas of difficulty.

The first step in the qualification process involves the entire faculty in the setting of program objectives. These objectives must be specific and must represent the best thinking of the faculty. Before objectives are adopted, the entire faculty reexamines them and makes any revisions necessary for the final product to be agreed to by all. To illustrate: Following this plan, the faculty members of a school of nursing in New York met to develop better clinical assessment procedures. They had originally planned to study several rating forms they had developed, but it quickly became obvious that some backtracking had to be done. The faculty had previously developed a statement of philosophy for the school of nursing. This statement included a list of eleven program objectives to be achieved by the time a student graduated. Following each of these objectives was a list of desired behaviors that varied from specific behavioral objectives to idealized attitudes that were almost impossible to assess with any degree of validity.

Working together, the faculty finally agreed upon two new broad objectives to be met by each graduate of the school.

In the second stage, these new broad objectives needed to be subdivided into a series of highly specific behavioral objectives. To do this, the faculty separated into work groups. Each group took the two broad objectives and applied them to each course taught within the school. For example, the group of faculty members teaching in the freshman year developed specific behavioral objectives for each course in the freshman year that led clearly and sequentially to the culminating goals of the school. These objectives were written in the form of behavioral criteria. Any item that in the judgment of the faculty could not be assessed objectively was eliminated. The remaining criteria became the equivalent of a curriculum guide for the faculty to follow in developing the clinical experiences. However, there was another important facet to this step.

Mandatory vs. Desirable Performance

We recognize that in any body of knowledge or in any profession there are some things one must know to establish a level of competence and other things that it would be desirable to know. The "must-know," or mandatory, material represents that specific knowledge or level of performance that the student must master if she is to be permitted to progress or graduate. To put it another way, this mandatory material represents the "pass" level--if the student does not know it or cannot demonstrate it, she fails. Our hope as educators would be to help each student go as far as possible beyond the pass, or mandatory, level.

To continue our example of the school of nursing: the faculty completed the preliminary steps of subdividing the broad objectives into specific course objectives and then further defined these course objectives in concrete behavioral terms so that all of them could be assessed objectively. At the same time, the faculty assigned to work at each level went through the list of behavioral objectives and skills and designated those that all agreed were mandatory, minimal levels of competency. To cite one illustration, the ability to measure the proper dosage of a medication was considered mandatory knowledge for safe practice. The knowledge of the history of the profession of nursing was considered desirable knowledge. If the student did not know the former, she failed; if she did not know the latter, she still went on. The mandatory items or areas formed the core of the planned classroom and clinical experience within each course. In addition, they were carefully explained to students at the beginning of each course along with the criteria and methods to be used to determine whether in fact a student had at least met the mandatory level.

There are several points to keep in mind in setting up such a qualification system. First, this system does not suggest that we should be satisfied with allowing students to slide by with minimal competency. By working cooperatively with each student towards the qualification goal, we concentrate on helping her learn and grow to her fullest capacity. We are more interested in each student's growth than we are in comparing her with her classmates. Second, we have found that the qualification system based on mandatory and desirable skills sharply reduces the assessment problems of the faculty and the evaluation threat to the students. Third, an essential part of the qualification system rests on an effective feedback system. This places on each faculty member the responsibility of constantly sharing with each student her progress towards meeting the stated qualification criteria.

Without going into great detail, suffice it to say that if a student reaches the end of a

course or clinical sequence and is surprised by her failure, the faculty has failed badly to provide the type of ongoing criticism the student needs in order to be aware of her progress at all times. There is no standard mandatory level of performance in nursing education. This must be determined by and agreed to by the total faculty within a given school. A good rule of thumb for determining true mandatory knowledge or skills is the following statement: Even if the student does well on everything else in the entire program, if she cannot do this or does not know this, she fails and will not be allowed to proceed. If the faculty agrees, then the item is truly mandatory. It should be emphasized that we are describing the final assessment checkpoint for a student after all attempts at remediation or tutorial assistance have been made.

Let us return for a moment and look at the process of feedback. To make a qualification system work, faculty must be able to help each other and the students, and the students must be able to help each other and the faculty to reach the objectives in the best way possible. Both faculty and students need to know what they are doing well, where they can improve, and whenever possible, ways in which they can improve or change their behavior.

We must, however, recognize that there frequently are barriers to giving and receiving help from someone else.

Barriers to receiving help:

1. It is hard to really admit our difficulties even to ourselves. It may be even harder to admit them to someone else. There are concerns sometimes as to whether we can really trust the other person, particularly in a work or other situation that might affect our standing. We may also be afraid of what the other person may think of us.

2. We may have struggled so hard to make ourselves independent persons that the thought of depending on another individual seems to violate something within us. Or we may all our lives have looked for someone on whom to be dependent and we try to repeat this pattern in our relationship with the helping person.

3. We may be looking for sympathy and support rather than for help in seeing our difficulty more clearly. We ourselves may have to change as well as others in the situation. When the helper tries to point out some of the ways in which we are contributing to the problem, we may stop listening. Solving a problem may mean uncovering some of the sides of ourselves that we have avoided or wished to avoid thinking about.

4. We may feel our problem is so unique no one could ever understand it, and certainly not an outsider.

Barriers to giving help:

1. Most of us like to give advice. Doing so suggests to us that we are competent and important. We easily get caught in a telling role without testing whether our advice is appropriate to the abilities, the fears, or the powers of the person we are trying to help.

2. If the person we are trying to help becomes defensive, we may try to argue or pressure him--meet resistance with more pressure and increase resistance. This is typical in argument.

3. We may confuse the relationship by responding to only one aspect of what we see in the other's problem and overpraising, avoiding recognition that the person being helped must see his own role and his own limitations as well.

There are certain qualities that serve to minimize or eliminate these barriers. They are:

1. Mutual trust.
2. Recognition that the helping situation is a joint exploration.
3. Listening, with the helper listening more than the individual receiving help.
4. Behavior by the helper that is calculated to make it easier for the individual receiving help to talk.

At the heart of the qualification system is the process of feedback. Feedback is a way of helping another person to consider changing his behavior. It is communication to a person (or a group) that gives that person information about how he is performing or how he affects others. The feedback process must involve all members of a school to be effective--faculty feedback to each other, student feedback to each other, and faculty-student feedback to each other. It is, ideally, a two-way communication: the receiver needs to know how he is perceived by the sender; the sender needs to know how he can more effectively communicate or help the receiver. The more open and honest we can each be with those around us, the easier it becomes for everyone to give more effective feedback.

Some useful criteria for feedback:

1. It is descriptive rather than evaluative. By describing one's own reaction, it leaves the individual free to use it or not to use it as he sees fit. By avoiding evaluative language--for example, using adverbs instead of adjectives--it reduces the need for the individual to react defensively.
2. It is specific rather than general. To be told that one is "dominating" will probably not be as useful as to be told that "just now, when we were deciding the issue, I felt you did not listen to what others said, and I felt forced to accept your arguments or face attack from you."
3. It takes into account the needs of both the receiver and the giver of feedback. Feedback can be destructive when it serves only our own needs and fails to consider the needs of the person on the receiving end.
4. It is directed toward behavior that the receiver can do something about. Frustration is only increased when a person is reminded of some shortcoming over which he has no control.
5. It is asked for, rather than imposed. Feedback is most useful when the receiver has evaluated himself and then formulates the kind of question that those observing him can answer. In other words, asking for feedback from others may be a way of also giving feedback. For example: "I get the feeling you are disinterested in what I am saying. Is there something I can do that might help you more?" If feedback is not asked for, it still must be given in order to insure that an individual has the knowledge of a deficiency and the opportunity to change. It may be desirable to first ask whether the individual would like to

receive feedback. For example: "I have noticed some specific things that seem to be interfering with your performance. I'd like to share them with you because you may find them useful." Or, "I notice that you haven't asked for feedback from me for a while. I wonder if there is something I do that makes it difficult for you to ask."

6. It is well-timed. In general, feedback is most useful at the earliest opportunity after the given behavior (depending, of course, on the person's readiness to hear it, on the support available from others, et cetera).

7. It is checked to insure clear communication. One way of doing this is to have the receiver try to rephrase the feedback he has received to see if it corresponds to what the sender had in mind. For example, asking the receiver to think about what has been said and briefly write out his understanding of it.

8. When feedback is given in a class or group meeting, both giver and receiver have opportunity to check the accuracy of the feedback with others in the group. Is this one person's impression, or an impression shared by others?

We are now ready to move to the next stage--that of creating an environment in which students can freely and openly demonstrate their growing ability and clinical skills. We are interested in each student's ability to apply the knowledge she has previously demonstrated in classroom situations. Without going into great detail here, let us assume that the faculty has also developed intermediary steps between the classroom and the clinical area. These would include role-playing, case studies, problem-solving exercises, simulation laboratories--even experimentation with the use of computer-controlled teaching systems to provide a simulated laboratory experience.

In order to plan appropriate learning activities in the clinical setting, the instructor is faced with several problems. One is the lack of a sufficient number of appropriate patients with whom students may work, and the other is how best to select patients with whom students may work. In a number of specialized areas, faculty members have difficulty in finding enough available patients with problems directly related to classroom content and clinical objectives. This relates closely to the other problem--that of the method used in selecting patients to be cared for by students. Proper patient selection becomes vitally important if students are to gain the proper experience under the previously described qualification system. Normally, the instructor selects the patients in each clinical sequence. However, this unilateral decision has been challenged in recent years. In general, if the learning process is truly a team effort between the faculty member and the student, then both should be involved in selecting and planning the appropriate learning experience in the clinical area.

Observation in the Clinical Area

It is important to mention several points at this time that have a bearing upon the effectiveness of learning and the assessment of progress in the clinical area. The more threatening the evaluation and observation process becomes, the less students are able to learn. For example, students frequently wonder whether or not they ever do anything right, for the only time they ever hear from an instructor is when they have made a mistake.

Perhaps the most common method of clinical supervision used at the present time is

the type of supervision provided by an instructor who, physically present, watches a student perform certain tasks. This has the advantage of having the instructor readily available if the student gets into difficulty and of enabling the instructor to be aware of everything the student does. It has the disadvantage of changing the dynamics of the situation. The student is aware of the presence of the instructor and may not react as she would normally. The patient frequently will try to support or protect the student if the instructor is present. Such supervision may be the safest approach, but it is of debatable effectiveness in observing communication skills or student-patient interaction. It also limits the amount of experience a student may have, because the instructor can observe only one or two students closely at a time.

A second method at least supplements the form of supervision described above. This method involves student self-reports of what occurred in the clinical experience. The reports may take several forms and operate under several names. The most common type of self-report deals with the nursing care the student gives a particular patient. One example of this is process recording. Other forms may be called interaction studies or clinical diaries. Still other variations of the self-report include student-written anecdotal records and nursing care plans or studies.

Both methods described above--physical supervision and self-reporting--represent the most common types of clinical supervision in use today. However, the use of audio-visual techniques as a supervisory and evaluative tool is beginning to increase rapidly. Nursing educators have long relied on the use of films, slides, and other similar techniques in classroom teaching. With the advent of portable video tape units, more is being done in utilizing video tape recordings as teaching and evaluative tools.

However, video taping has not yet been widely used in the clinical area. The method is almost standard practice in many other helping professions. There are several ways in which it can be used in nursing education. A portable camera and a video tape recorder could be used in home visits or in selected clinical situations, always with the patient's permission. Those schools of nursing that have access to a clinical unit that monitors individual patients by closed-circuit television can hook a video tape recorder into the master control unit. Although initially costly in terms of equipment ($1,500 to $2,000 per unit would be the current range), this method has the greatest promise for clinical teaching and supervision. An instructor who is able to view the video tape of a student caring for a patient is able to point out to that student both verbal and nonverbal behavior more effectively than ever before. The student can also view the tape with classmates as an additional reinforcement to learning. For those schools without the facilities or the funds to utilize video tape techniques, there is one alternative available for clinical supervision--the use of audio tapes.

Instructor Reports

Up until this point, I have been discussing primarily self-report techniques and methods of gathering information about student progress and learning. However, just as learning is a two-way process between faculty and students, so is the gathering of information. As faculty members, you are charged with the responsibility of ultimately determining a student's ability to function in a competent, safe manner in the clinical area. In order to do this, you need all possible information. The information may be gathered and reported in several ways. One way includes anecdotal material based on systematic observation and structured interviews. The other way includes written

methods that are more structured, including rating scales and checklists.

Anecdotal records are among the data-collecting devices used most frequently in the observation and evaluation of student behavior. The anecdotal record can best be used in reporting conversation and interpersonal relationship, where variations in situations and personalities preclude any advance determination of behavior. Anecdotal material can also be employed to show trends in a student's behavior by presenting specific brief incidents related to her use of nursing skills and techniques or to her use of judgment. Since the instructor's opinion is kept separate from the anecdote, anyone else reading the account of the incident is free to form her own opinion on the basis of the facts presented.

The anecdotal record lends itself well to frequent observations of short duration. One anecdote by itself is of little or no value, but a series of notes can be very helpful in pointing up persistent behaviors or trends over a period of time. Ideally, there should be a number of anecdotes written by different instructors about each student over a period of time. Using more than one source tends to lessen individual bias. The clerical burden of maintaining good anecdotal records can be greatly reduced by providing dictating equipment; anecdotes can later be transcribed by a secretary.

Anecdotal records should be made available for review by students at any time. They should be periodically summarized and individual reports should be destroyed. A good inservice training program consists of having a faculty member write some anecdotes, share them with other faculty members, and ask for specific criticism. The instructor should be looking particularly at the way the anecdote is interpreted by each reader. If there are sharp differences of opinion, then the anecdotal record is probably too ambiguous or too subjective.

Perhaps the most common method currently in use in clinical evaluation for arriving at a grade is the rating scale. The scale has two main features: (1) a description of the characteristics to be rated and (2) some means by which the instructor may indicate the quality, frequency, or importance of each item. A good rating scale is built according to the following guidelines:

1. It is important to determine specifically what you are trying to find out.
2. Faculty must then determine, on the basis of clinical objectives, the most effective ways of measuring those things that can be measured.
3. The scale should contain clearly defined traits, expressed in the form of brief descriptions of behavior that can be observed.
4. Space should be provided on the scale for the faculty member to record instances or anecdotes that will support the rating given.
5. When the faculty member has had no opportunity to observe the specific behavioral item to be rated, she should be encouraged to indicate this fact on the form.
6. The scale should be limited to behavioral items chosen because of their importance.
7. The scale should be accompanied with specific directions for completing the form.
8. Both the scale and the instructions for completion should be constructed in cooperation with those who will use the form.

9. It is better to have four categories than five because then, faculty members are forced to discriminate by placing students in other than a middle, or average, category.

The building of a good scale takes time and effort. A series of brief scales is preferred for measuring specific skills; broader categories can be used to measure behavior that is common to all clinical situations.

Although much has been written about the development of rating scales, it is entirely possible that such scales may be completely unnecessary for clinical evaluation. Previously, I described a qualification system based on mandatory and desirable knowledge. If a school is using this type of satisfactory-unsatisfactory or pass-fail system, there is little need for rating scales. Instead, simple checklists of behaviors and skills broken down into mandatory and desired behaviors will serve the purpose of evaluation without the problems inherent in building, validating, and using rating scales.

The success or failure of a good checklist rests on the grading system used in the clinical area. If the faculty members of a school can reach consensus on clinical objectives, behavioral criteria, a grading system based on pass-fail or S-U, and behavior that is considered mandatory, then one or more checklists may solve their problems of evaluation in the clinical area. When a student has demonstrated at least minimal mastery of a particular skill or technique,the instructor checks off the appropriate item on the list. In order to pass, the student must satisfactorily complete all mandatory items in the clinical sequence. Ideally, the student would be checked for each item or skill in different situations by different instructors, but this may not always be possible. However, a safeguard can be built in by requesting that another instructor observe a potentially failing student as a validity check. Checklists have been found to work very effectively in highly specialized clinical areas. With faculty involvement and agreement on what is to be emphasized, there seems to be no reason why checklists could not be expanded to cover all clinical situations. It is important that no standardized checklist or critical-incident form be used. Each checklist must reflect the specific clinical emphases of a specific group of faculty members if it is to work effectively.

Self-ratings and Peer Ratings

In describing the use of rating scales and checklists, I have concentrated on forms and procedures for assessment to be utilized by instructors in the clinical area. However, there are other supplementary methods that can greatly enhance the evaluation process. It should be remembered that we have been emphasizing a system of evaluation that involves both the student and the instructor in the learning process. The same principle applies in the area of assessment and evaluation--that is, the student should be an active member of the evaluation team.

I have suggested some guidelines for the development and the utilization of rating scales and checklists based on clearly stated behavioral objectives. I have stressed the fact that objectives and criteria for evaluation must be clearly communicated to students at the beginning of each clinical sequence and that individuals should be given constant, detailed feedback with regard to progress towards meeting the qualification of successful completion criteria. At the end of each clinical sequence and at periodic points during the sequence,instructors complete evaluation forms for each student. At the same time, students can be asked to complete the same form, rating themselves, independently of the instructor.

At the clinical, or final, evaluation conference, the evaluation form completed by the student can be discussed and compared with that completed by the instructor. It may be that both agree on current progress and future plans. However, discrepancies need to be examined closely. If the instructor rates the student high or passing in a particular area and the student rates herself low or failing, then the instructor may need to deal with a student who has a very poor self-concept or an unrealistic estimate of her own ability. On the other hand, if the student rates herself high and the instructor rates her low, it may indicate that the student's self-appraisal is very unrealistic in the face of specific evidence. However, it may also be possible that the student is correct in her self-appraisal but is giving the wrong signals in her methods of communicating with others. Or, it may be, too, that she really can function in the particular area and should be given another opportunity to demonstrate the required skills.

In other words, the dual assessment by the student and the instructor, with each open to change in the light of the evaluation conference, can greatly facilitate the effectiveness of the assessment process and the final evaluation conference. It serves to provide both with more information, a result that inevitably will increase the validity of clinical evaluation. The same principle applies to the use of peer ratings. These will be effective only when completed anonymously and with a high level of trust between students and faculty. Systematic peer feedback can be given in a variety of ways, both written and oral. When used properly, it helps to add a third dimension to the evaluation process. I recognize that the final responsibility for evaluation must rest with the faculty. However, each individual student has a responsibility to herself, her peers, and the profession she is entering. Therefore, she should share in the responsibility for systematic, open, and honest assessment of her own strengths and weaknesses and those of her classmates with whom she works. Ideally, this use of self-ratings and peer ratings could also greatly enhance the effectiveness of any health-services team working regularly together.

Evaluation Conferences

The progress reports, or final evaluation conferences, are a vitally important part of successful and educationally sound clinical evaluation. These conferences should be carefully planned for by both the student and the instructor. Both must have a clear understanding of the purpose of the conference in advance. Ideally, it is to provide a time for the sharing of information about the student's progress in as nonthreatening an atmosphere as possible. Both should have the opportunity to gather whatever information it is necessary to bring to the conference.

In 1965, the results of a faculty debate (see Bibliography) on the pros and cons of individual clinical conferences were published. On the negative side, it was argued that too much time was invested and that it was unpleasant to debate grades with students. The majority concluded that if time permitted, frequent individual conferences should be held with students. In many respects, this was a poor debate with a poor conclusion. It is not a question of whether or not time permits. It is essential to student-faculty communication that an open, valid, fair, and consistent evaluation system be developed, maintained, and safeguarded. Without regular individual conferences, evaluation tends to become biased and autocratic, for the instructor never needs to defend her evaluations. In other words, there is no accountability on the part of the instructor for her actions. Many of the desirable elements of a good conference were summarized by Sister Dolores

Calamari in 1968 (see Bibliography). Both students and faculty can positively anticipate evaluation conferences if the guidelines suggested in the 1968 report are followed.

Summary

The points I have tried to share with you this morning will not necessarily solve all problems in clinical evaluation. The professional judgment of the faculty is and always will be important. However, we need to constantly reexamine what we do and how we do it. We especially need to involve students much more than ever before in decisions affecting their destiny. I would have students on all committees within the school, especially on the guidance and evaluation committees. They need both greater freedom and greater autonomy, as well as the responsibility that goes with them. If we offer our students the same trust, the same commitment, and the same support we offer our patients, then we truly join the ranks of educators in a helping profession.

Bibliography

Calamari, Sr. Dolores. "Factors That Influence Evaluation Conferences in Clinical Experience." J. Nurs. Educ. VII:4:11-14, Nov. 1968.

Cobb, M. Marguerite, Rose, Patricia A., and Schumann, Delores. "The Pros and Cons of Formal Evaluation Conferences--A Faculty Debate." Nurs. Forum IV: 4:57-75, 1965.

Litwack, Lawrence, Sakata, Robert, and Wykle, May. Counseling, Evaluation, and Personnel Services in Nursing Education. To be published by W. B. Saunders, Philadelphia, in early 1972. The book presents in greatly expanded form many of the ideas included in this paper.

EVALUATION--AN OBJECTIVE APPROACH

Presentation by Lawrence M. DeRidder

Mildred Montag, in her foreword to Alice Rines' monograph on Evaluating Student Progress in Learning the Practice of Nursing, stated, "While we will never be able to perfect evaluation practices, we can and must develop better ways and means of evaluating the student's performance. It is not an easy task." (1963, Foreword) In a similar fashion, Stake approached the problems and the process of evaluation thusly: "Evaluation requires judgment. Decision-making requires judgment. Both are judgmental in themselves but also depend on judgments previously made. A school and a curriculum are where they are because of judgments made from within and from without. Judgments are made early, and late, and in-between times. To understand what a school is doing requires an understanding of what a school is expected to do. In education as elsewhere, judgments will continue to rest on incomplete knowledge, imprecise measurements, and inadequate experience. No error-free system is possible, but improvements are within easy reach." (1970, p. 181)

You and I are involved in education because of our own personal value commitments, our educational objectives, our priorities, our perceived norms and standards. Each of us, directly and indirectly, influences what happens to our schools. One way to determine the degree of consensus among us is to examine our objectives. The stated objectives are generally concerned with a change in behavior such as a changed attitude, a higher perception or skill level, or an increase in knowledge. Whatever we list, however, is always an oversimplification, for our choices--the result of our combined judgment--are the more valued goals of the many that are possible. Combined judgments, of course, achieve greater acceptance than individual judgments, which is as it should be.

Objectives, then, really originate in the values of the group, and agreement on certain ones indicates that we give them priority. The next step is to rank the chosen objectives in order of desired emphasis.

What qualities should our instructional objectives have? Lindvall and Cox (1970) suggest the following:

1. Instructional objectives must reflect our agreement on the content of our teaching and testing.
2. They must be free of ambiguities in order that they can be used by lesson writers, teachers, and evaluators.
3. They must be placed in a sequence determined by levels of student mastery and progress.
4. Well-written instructional objectives permit lesson writers to develop sequences of nonoverlapping or missing lessons aimed at achieving successful progress of students.

Lawrence M. DeRidder, Ph.D., is Professor and Head, Department of Educational Psychology and Guidance, The University of Tennessee, Knoxville, Tennessee.

5. They should show no evidence of important gaps in learning.

Such objectives fit into my three-part definition of evaluation: (1) the process of determining the kinds of decisions that have to be made, (2) the systematic attempt to select, collect, and analyze evidence concerning those changes in student behavior that accompany planned educational experiences, and (3) the use of this information to make decisions about the educational program. The goals of the program, in the form of behavioral objectives, define the areas of decision-making--the end products of the learning experience. An instructional program is developed, level by level and course by course, for the achievement of each objective and, finally, an assessment of the achievement of the program's goals is made and fed back into the system.

It may be helpful at this point to review briefly some principles of evaluation identified by Sorenson. The first principle he identifies is one the Council of Diploma Programs apparently has already agreed upon: the purpose of any instructional program is to produce measurable changes in the students for whom it was designed. It follows, then, that the goals of our instructional program--whether they involve knowledge, feelings, or attitudes--must be defined in terms of performance, behavior, or actions. Instructional procedures are then designed to meet these prestated goals. As we develop our instructional procedures, we must agree upon how to arrange sequences of learning tasks, when and how to provide learning incentives, when and how to give students feedback on their performances, et cetera. When possible, each part of the instructional program should be tried out with a small sample of the student population to determine what changes in objectives or instruction should be made. It is important to know the reasons why some students did not learn. Could their previous learning have influenced them unfavorably, for example? Or were they not motivated? Only when we know why they failed can we develop appropriate instructional programs for them.

A continuing evaluation program is necessary to examine the effects of the instructional program and to determine where revisions are needed. And so it is, as Rines (1963) has indicated, our objectives that set the standards against which the student should be evaluated. However, these standards must be relevant and accurate, or they have no worth; our objectives need to be checked against the judgment of our experts: nurses on the job, professional nurse educators, and students in training. This procedure is predicated on the assumptions that we can agree on our nursing goals, that we can order the objectives both psychologically and logically into a pattern or sequence that will permit successful attainment of these goals, that we know how successful learning can best be encouraged and accomplished, and that we can identify evaluational approaches that will provide us with data that can be used to assess the success of our instructional program.

What I have already said suggests some of the purposes and end products of a well-designed and well-executed evaluational procedure based upon behavioral objectives.

1. A primary purpose is to determine the adequacy of the instructional program and of each of its components in producing successful nurses.

2. By defining and clarifying our educational objectives, we can identify what can easily be eliminated and what objectives require changes.

3. The course-by-course and level-by-level evaluation helps to make us aware of each student's progress toward achieving the goals.

4. Because we know what we want to achieve, we can determine more easily how ready students are for the unit and, on the basis of student limitations or weakness, modify our instructional materials and techniques to provide successful learning opportunities for a greater number of students. If a student does not learn in class, the defect is likely to be in the instructional procedures. The net effect is that the teacher is helped to improve her teaching. As Dressel and Mayhew stated, "The instructor who critically regards his practices every day, who seeks to find in his students evidence of the effect of his instruction and of their own endeavors in terms of progress toward accepted educational goals, will find that evaluative-centered thinking is a continuing, rewarding, and integral part of his teaching." (1954, p. 24)

5. All students will learn a great deal more, regardless of I.Q., aptitude, or limited cultural background.

6. Because each student knows in advance the competencies the unit of learning is expected to develop and because she has an opportunity to get an early assessment of her current functioning level, she is more highly motivated and learns more rapidly.

7. Thus, both students and staff have greater psychological security--they know where they are, where they are going, how they are going to get there, and when they have arrived. Additionally, the program readily permits assessment of the competencies of transfer students in order to identify appropriate course and level placement. New staff are also more easily oriented to their teaching responsibilities.

8. The instructional program is developed and tailored for acceptance by particular teachers who will use them effectively because they were involved in the total process; additionally, the program is designed with sufficient flexibility to fit the total student population of a specific school of nursing. Effective teachers will vary their materials and techniques with different classes and even with different students.

9. A wider range of reliable and valid instruments for evaluation are identified and used.

10. Students and staff are more comfortable with and acceptant of the final course and program evaluation, since the whole process is understood.

11. The legal requirements of nursing certification are met.

Assuming that we have developed acceptable classroom and clinical lab objectives and that these objectives have been both grouped logically and sequenced according to level of learning and skill complexity, how do we actually evaluate student accomplishment? How can we know when the desired learning has occurred? According to Bloom, ". . . the clearer the specifications are in terms of both conduct and behavior, the better. Such specifications define the problems they must solve in the construction of instructional materials or evaluation instruments and such specifications provide the criteria against which the materials and instruments are validated." (1969, p. 28)

The specifications of the objective suggest the types of evaluation procedures to be used: Clinical objectives may require observation of the student or observations made

by the student; academic or classroom-oriented objectives may possibly involve cognitive tasks or affective responses of the students, open- or closed-book exams, essay or multiple-choice questions, et cetera; both clinical and classroom objectives might call for frequent testing or comprehensive examinations as well as examination of products and processes. The specifications of the objectives, both clinical and academic, will determine the weight and importance to be given to particular aspects of the evaluation, and even the performance standards to be used.

A primary task of teachers (and evaluators) is to design the evaluative process so that it has a positive effect on student learning and instruction and leaves both teachers and students with a positive view of themselves, of the subject, and of the learning process. In addition to traditional tests, many other methods of obtaining evidence of changed behavior must be explored.

Some necessary conditions are:

1. The data-gathering measures must be valid: they must examine, by the most direct methods, those learning outcomes regarded by teachers and students as important and desirable. Evaluation can be no better than the internal validity of the information presented to it. If observations are utilized, the observers need to be so trained that the data sought is equally perceived and recorded by all observers.

2. The data gathered must be regarded as reliable and objective: chance and error must play a minimal part in determining the adequacy of each examinee's performance. If paper-and-pencil tests are used, the items should not depend on special language learned in the course unless the learning of the language is part of the objectives.

3. Because evaluators bring to evaluation their own needs, feelings, and biases, it is important that the teacher understand how her perceptions can influence evaluation. The use of illustration and the sharing of what ratings mean to different members of a group assist in clarifying the use of ratings and improves reliability. In the St. Mary's study, selected anecdotes were rated and discussed by the staff in order to achieve group consensus regarding rating levels and internal consistency in the rating of observed behavior. The standards for grades must be defined in terms of adequacy rather than of rank order of students and competition--that is, a student is judged in terms of what she has been able to do. The objectives and their specifications determine the required quality of performance. Since nursing necessitates both competence in clinical performance and a specific academic background, it is quite conceivable that a student might succeed in one and do poorly in the other. In this case, she would be assisted in the area of deficiency until she has demonstrated the desired standard.

The problem of evaluating performance has always been a difficult one. Several choices are available to each of us. After an agreed-upon performance criterion has been established, we can use existing appraisal systems or rating forms or select relevant sections from them or modify them to meet the specific purposes of the evaluation, or we can make our own.

If we are ingenious, we can usually figure out some way to acquire objective data for assessing effectiveness. Dr. Mildred Katzell (1967), in her presentation "Evaluation:

Its Meaning and Use," suggests, as an example, the use by teachers of different-colored sheets on which observed behaviors are recorded. When a student is observed doing an important aspect of her job very well, it is written on a green sheet and filed. When an ineffective performance is observed, it, too, is recorded, perhaps on a yellow sheet, and filed. In this way, the observer becomes a describer of behavior rather than a judge. When the instructor reviews the data with the student nurse, she has an opportunity to examine the reasons for both effective and ineffective behavior specimens and can, perhaps, find ways of being more helpful. An alternative approach, similar to the "critical incidents" approach, might be writing brief statements describing an outstanding, a satisfactory, and an unsatisfactory performance of a particular skill.

Most of us are already familiar with a large number of different methods of obtaining evidence of behavioral change. Among the most useful are:

1. Tests of aptitude, ability, and achievement.
2. Self-inventories concerned with interests, attitudes, adjustment, appreciations, temperament, semantic meanings of concepts.
3. Rating scales, achievement checklists, checklists of activities.
4. Teacher-made tests of knowledge, skill, problem-solving, and attitudes.
5. Records of direct observations, such as anecdotes.
6. For some objectives, unobtrusive rather than direct methods of assessing change may be highly appropriate--for example, reference to changes in grade point average, awards, absences, days in attendance, withdrawal, transfers, referrals, telephone calls, books and magazines read, peer nominations, elective offices held, frequency and types of disciplinary actions, requests for information, conferences held, concerts attended, membership in professional or social organizations, time spent in community activities, et cetera.
7. We now have techniques to measure (a) the institutional characteristics, such as the Environmental Assessment Techniques (Astin); (b) the perceptions that students have of their environment, such as the College Characteristics Index (Pace and Stern) and the College and University Environment Scales (Pace); and (c) the observable student behaviors, such as the Inventory of College Activities (Astin).

An inevitable and continuing headache for all teachers is how to grade or evaluate student performance. The evaluation model presented places emphasis throughout on the extent to which behavior changed as a function of good instruction. The emphasis is not upon the learners but upon learning and, consequently, upon the effectiveness of instruction. Measurement of changes in behavior produced by instruction are known in the literature as formative evaluation. For this type of appraisal,we identify anecdotes, observations, or test items that are directly related to the behavioral objective as opposed to items that are able to discriminate among students. A clear statement of the objective will provide the performance levels needed and suggest appraisal techniques for measuring the objective. The clinical-practice objectives as well as the subject-matter objectives would be equally examined. Any approach used to evaluate would have content validity only if it adequately represented the content to be mastered. By definition, then, formative appraisal items that evaluate learning resulting from instruction would, generally,

be difficult for learners before instruction and easy after instruction. If they were missed by many students, this would suggest the need to modify instructional procedures, not that the students should be given a low grade. The primary concern is to examine the extent to which each student has achieved the criterion or specifications of the stated objective, not how well she is doing as compared to other students. Performance standards that have been previously specified and communicated to students are used as a basis of assessing student growth and in determining whether an educational program and all of its component parts have achieved their purposes.

Brief criterion-oriented appraisals, when given frequently, are helpful in assisting students to pace their learning, in determining their competence in each task of a sequence, and in motivating them. Each student should have a copy of the objectives for each unit and know well in advance the kind of appraisal approach the instructor will use. The appraisal results also provide the instructor with feedback of errors in instruction, which serves as a corrective in the teaching process. In addition, they provide the students with feedback concerning their learning, informing them of what they have mastered and where they are having difficulty. Such feedback is most useful when it not only identifies what the student must yet learn but also suggests specific instructional materials and procedures she should use in order to bridge the gaps in her learning. Remediation, if the required level of learning is not achieved, would be both natural and normal. The student has a chance, through the help of the teacher, to produce successfully. On the other hand, the objective itself or the instructional techniques of the teacher may need to be examined.

Preferably, these partial or formative evaluations should be separated from the grading process and used primarily as aids in the teaching-learning process, although they may be graded in order to give the student some idea of her current competence. However, whatever might be her initial learning limitations, either in class or "on the floor," no student is penalized in her final evaluation for an inadequate background or for errors in learning. In other words, grades given on each unit are not averaged with end-of-course tests to get a final grade. The evaluation used at the end of the course, term, or educational program (the summative evaluation) should represent the individual's efforts in achieving the objectives.

Because students differ in their performance even with the same instruction, some evaluators suggest that appraisal procedures be written at different levels of difficulty to cover the range of possible performance as related to the objective. Moreover, if the behavioral objectives have been ordered according to levels of complexity, appraisal methods could be included that measure performances that should have already been achieved as well as observations or test items that measure advanced achievement. This would help to determine any unanticipated outcomes, the extent to which the student has met or surpassed the competency level of the objective, and the direction in which future instruction should go. If more students achieved a complex performance and did less well on items that are more basic, it might well mean that the sequencing of objectives needs to be restudied.

Although many items should be included in an appraisal instrument in order to cover the full range of the performances desired, the items relating to specific objectives should be grouped and scored separately in order to provide both the student and the teacher with diagnostic feedback. Appraisal instruments so constructed can be used by teachers to make decisions about pacing, review, or termination of instruction, and by students to determine their progress and as actual study materials.

The teacher-learning approach outlined thus far is performance-based or criterion-oriented. The emphasis is on student mastery of the specified objectives and the content of instruction. Under this system, it would be entirely possible for all students to attain mastery and, consequently, obtain an A grade. Of course, it is equally possible that few or none of the students obtain an A grade.

Determining at which level of performance different grades are to be given remains a problem. Some of us who have moved in the direction of grading on the basis of mastery have applied standards determined for similar tests or appraisal instruments used earlier with similar students. Students are then informed that they will be assigned grades according to these predetermined performance levels, not according to their rank order in the group. If you follow this approach, you should set a realistic level of performance attainable by most students in your school.

Some spin-off benefits of this system are likely to be already apparent to you:

1. Students are encouraged to work with and help, rather than compete with, each other.

2. The communication between the teacher and the student is enhanced, since the teacher fits more easily into the role of facilitator of learning.

3. Since research results suggest that, given enough time, most, if not all, students can achieve the mastery performance, nurse educators can look for ways to provide each student with the time she needs to learn a subject. The time needed may vary greatly among the students and depends upon such variables as verbal ability, ability to read, scholastic aptitude, prior achievement, the quality of instruction received in class, the amount and quality of help received out of class, and the interest level of the subject matter.

4. Tutors can be provided for those who encounter learning hurdles.

5. Students may be permitted to move through the course or program at their own speed. Some of my students are able to display such a level of mastery on the final examination given during the first day of class that they are encouraged to move on to more complex and challenging learning material.

6. The teacher has more rapid feedback on the effect of instruction and, consequently, is encouraged to explore some of the newer and frequently more effective ways of achieving successful learning: programmed material, simulation, role-playing, computer-assisted instruction, audio-visual materials, independent research and study, et cetera. As suggested earlier, the use of brief diagnostic progress appraisals are very helpful in determining which students have mastered a clinical-practice or academic unit and what the other students must yet do to master it. Such appraisals help to insure mastery before a new and more difficult learning task is attempted. Students who need help can be referred to the specific instructional materials intended to correct the identified learning difficulties.

7. Students who need specific assistance may be grouped together to review and help each other to overcome their common learning problems.

8. Most importantly, the student, when informed that she has mastered what is required, gains in self-confidence. As a result, she begins to enjoy the sub-

ject and wants more of it. Frequent reassurance of progress builds a more positive self-concept and, consequently, better mental health.

Controversy over what is the most appropriate grading system continues to rage, within academic circles particularly. Most schools have already rejected systems of grading requiring more than the five-point system and are moving to more simple and possibly more valid procedures, such as honors-pass-fail or just pass-fail. Miller summarized the arguments against the current A-F grading system as follows:

1. Conventional A-F letter grading corrupts the teaching situation by encouraging "grade-grubbing" and alienating students from the teacher.
2. It is neither accurate, being dependent on too many nonacademic variables, nor is it uniform. A grade from one instructor or one college does not mean the same thing as a grade from another.
3. It encourages a mechanical kind of learning and restricts the creative leisure of the student. It encourages conformity in some students, cynicism in some, and rebellion in others. It is certainly out of tune with the aspirations of the new student generation.
4. It does not, as is claimed, significantly stabilize the academic standards . . .
5. There is little evidence that it motivates students, less that it motivates them to learn significantly. What evidence there is suggests that learning proceeds more efficiently when students are encouraged to motivate themselves.
6. It serves the purpose of determining graduate admission, but badly. . . . There is evidence that the graduate schools could do without conventional letter grades.
7. There is evidence that transfers could be handled without conventional grades.
8. There is evidence that businesses can manage to hire without conventional grades.
9. There is a great deal of evidence to show that the grading system has no counterpart in real life, that grades are not related to success except in academically allied fields, and that it is in many ways a negative preparation for life.
10. There is considerable evidence that the grading system discourages and finally weeds out the most creative students, that it discourages creativity in general.
11. Grading protects and even encourages bad teaching. (1967, pp. 21-22)

What Miller did not say was that A-F grading is usually the result of such norm-referenced tests as standardized tests of achievement, intelligence, and aptitude, which measure an individual's performance in comparison with that of other students. Such instruments are very effective in ranking individuals or groups--not only in the same school but in different schools as well--in terms of their ability or knowledge. They have often served, too, as a basis for decisions concerning selection, promotion, and even program evaluation. However, they supply little information regarding specific skills and knowledge already acquired. They do describe individual differences, but not

for the purposes of instruction. The typical standardized test and some teacher-made tests are norm-referenced and do not serve as efficient sources of information, for which they are used and relied upon. This is so because in norm-referenced instruments, we retain only the test items that discriminate among students and we are likely to omit content-valid items that students may have learned as a result of our instruction. Such tests usually do not have any visible instructional objectives, or if instructional objectives are stated, the relationships of the objectives to the test items and the process of assessing instruction are vague. By concentrating upon finding differences among students, we neglect finding out how much the student has learned and what the roles of the environment and the intellectual processes are in producing the learning. The results of norm-referenced tests do not suggest ways of revising instructional objectives or methods.

As Bloom (1968) has indicated, we have become so used to the normal curve in grading students that we have an unquestioned faith in it. On any test, no matter how adequately or inadequately it measures what it should measure, we distribute grades on the basis of the rank order of the students in the group, so that a few get A's and a few F's. Consequently, some fail even though they may have achieved mastery of what was taught. This approach assumes that only a few of our students should learn. Administrators and even a teacher's colleagues frequently get "up tight" if her grade distribution is other than normal. What is worse, we label students through grades: "She is a D student," for example. Bloom suggests instead that "If we are effective in our instruction, the distribution of achievement should be very different from the normal curve . . . our educational efforts have been unsuccessful to the extent to which our distribution of achievement approximates the normal distribution." (1968, pp. 2-3) In other words, Bloom is supporting the need for content-valid, rather than norm-referenced, tests if we wish to make decisions as to the kind of instruction to provide or to determine the effectiveness of a program in achieving its objectives.

What are some alternatives? Unfortunately, there is no one solution. Each school will have to find a particular evaluation system that is compatible with its own educational objectives. Among the alternatives, the pass-fail grading approach is currently being tried out by many schools. Yale University, for example, put all of its undergraduate courses on a pass-fail basis in 1967. Some of the objections to it are as follows:

1. Grades still exist; the teacher still evaluates and assigns a P or an F.
2. Instructors are likely to vary in assigning grades and the grades themselves will suffer from inaccuracy.
3. Identifying students for scholarship and other honors would be impossible.

Some possible values are:

1. Evaluation will be less variable among teachers.
2. Nonacademic biases that affect grades will be less influential.
3. A poor teacher will have a harder time keeping learners captive.
4. Students are more likely to be self-motivated and find more appropriate ways than grade-getting to influence their peers.
5. Cheating will diminish.
6. Students who wish academic distinctions would be encouraged to experiment and to engage in creative projects.

7. The teacher will be free to encourage different approaches to learning and achievement.

A variant of the pass-fail is the pass-withdraw, or A, B, C, or W. Such an approach is currently in use at the University of Rochester. This system permits a student who feels he is prepared in a particular area to take the final examination during either or both of two examination periods during the year. On the basis of preset standards, grades of A, B, C, and failure are announced, but only passing grades are recorded. A student may take an examination as many times as he likes and may choose which grade among those he receives he wishes to count toward his program. A detailed syllabus, with suggested readings, is periodically published for student use. Each examination is constructed to measure the competence acquired by the student in two or more courses. Although grades may be given in courses, they are only for the information of the student, since degree requirements are based only on accomplishment as determined by the examinations. A modification of the above is now being tried in several courses in the Department of Educational Psychology and Guidance at The University of Tennessee (Knoxville). Students in these courses are not permitted to fail, but may withdraw if they wish. Those who appear to be succeeding but need more time to meet the preset standards may request an "Incomplete."

The obvious advantages include:

1. The emphasis is on success rather than on failure.
2. The student is given an opportunity to request a time extension when it is needed to achieve the mastery level desired.
3. The teacher is afforded flexibility in determining the instructional input.
4. Pressure upon the student to perform within relatively narrow time and, possibly, content limits is removed.
5. Both student and instructor have an opportunity to keep communication and, consequently, instruction open until the level of competence that is mutually acceptable as well as passing is reached.
6. The instructor exercises freedom and responsibility in determining objectives and in making them known and available at the time of the first class meeting.
7. Pressure is put upon the instructor to plan, prepare, and produce, since students are no longer a captive audience and are not likely to remain in an unpleasant or uninteresting course.

Some limitations may also emerge:

1. Students may give the courses graded with A-F their first priority of time and effort.
2. Grades, in the form of a P, are still given; evaluation is still necessary.
3. Some variability among instructors regarding the meaning of P may still exist.
4. The problem of interpreting P to other schools and to potential employers, while no greater than that involving A-F, leaves no opportunity to indicate the adequacy level of performance.

Still another approach is to move to comprehensive or qualifying examinations, which may coexist with or without pass-fail or conventional A-F grades. In this approach, grades given or not given prior to the comprehensives have no effect upon potential completion, since the student's competence is assessed through summative examinations based on the program's objectives. The University of Chicago has used this approach.

Some objections may be the following:

1. Instead of trauma over grades course by course, the trauma is focused on one final exam.
2. Too many or not enough diverse routes to final achievement may exist.

The advantages some have found in this system include:

1. Time, energy, and expertise might be found to develop highly reliable and valid examinations.
2. Students no longer compete but rather cooperate in preparing for exams.
3. Instructors have to define in advance the objectives of the program--what the student is expected to learn.
4. Students are examined for mastery of skills and subject matter, not to determine their relative standing.
5. Students may obtain an exhaustive analysis of the exam results, which assists them to identify their strengths and limitations.
6. Students are informed in advance of what they are expected to learn and are provided with the resources for learning: courses, syllabi, reading list, practice exams, et cetera. Also, they could either work independently, or with others in small groups, or take regular courses.

What has been presented has taken too long. That is really one of the major problems with evaluation--it is never-ending; there is no sure way; it is essentially a matter of expert judgment and decision-making applied to behaviors that never stand still. To the extent to which we can identify the behaviors we need, produce them, and measure the change we have produced, to that extent we have met the enemy and he is ours. We as teachers must be held accountable for the products we produce, which accountability includes the learning conditions with which we provide the students.

The approach I have outlined will take a lot of doing and undoubtedly involve a few false starts, but the end result, more successful learning for more nurses, will be well worth the effort.

Bibliography

Bloom, Benjamin S. "Learning for Mastery." Evaluation Comment I & II:2-3, May 1968. Los Angeles, Center for the Study of Evaluation.

_____ "Some Theoretical Issues Relating to Educational Evaluation." In R. W. Tyler, ed., Educational Evaluation: New Roles, New Means. 68th Yearbook, Part II. Chicago, National Society for the Study of Education, 1969, p. 28.

Dressel, Paul L., and Mayhew, L. B. General Education: Explorations in Evaluation. Washington, American Council on Education, 1954, p. 24.

Katzell, Mildred. "Evaluation: Its Meaning and Use." Unpublished speech presented at the Nechen Conference, St. Anselm's College, Manchester, N. H., Oct. 12, 1967.

Lindvall, C. M., and Cox, R. C. Evaluation as a Tool in Curriculum Development: The IPIE Evaluation Program. Chicago, Rand McNally, 1970.

Miller, Stuart. Measure, Number and Weight: A Polemical Statement of the College Grading Problem. Knoxville, Tenn., Learning Resources Center, Sept. 1967.

Montag, Mildred. "Foreword." In Rines, Alice R., Evaluating Student Progress in Learning the Practice of Nursing. New York, Teachers College, Columbia University, 1963.

Rines, Alice R. Evaluating Student Progress in Learning the Practice of Nursing. New York, Teachers College, Columbia University, 1963, p. 3.

Sorenson, Garth. "Evaluation for the Improvement of Instructional Programs: Some Practical Steps." Evaluation Comment 2:4:13-17, Jan. 1971.

Stake, Robert E. "Objectives, Priorities, and Other Judgment Data." Review of Educational Research 40:2, Apr. 1970.

EVALUATION--AN OBJECTIVE APPROACH

Presentation by Mildred E. Katzell

The title of this presentation was assigned to me, and I suspect the pun involving objectives was entirely deliberate.

It may be a surprise to you to learn that there are respected and respectable authorities who question the use of behavioral objectives in education, among them Robert Ebel, who is prominent in the field of educational measurement. Nevertheless, most definitions of evaluation in the books today relate evaluation to objectives. The definition that I prefer is found in Norman Gronlund's Measurement and Evaluation in Teaching: "Evaluation is a systematic process of determining the extent to which specific objectives are attained." It is clear and concise, and contains the essential ingredients. It says that evaluation is systematic; not every observation or judgment is evaluation. Evaluation is also a process--a series of actions, not a single isolated operation. It says that our objectives must be specified, so that we can determine the extent to which they have been attained. It acknowledges that our objectives may not actually be attained. It permits both quantitative and qualitative descriptions of student performance, combined with teachers' judgments of the value or desirability of the performance. It recognizes that evaluation cannot be totally objective, since value judgments must be involved.

As we set out across the wilderness of evaluation, we must keep our purpose in mind: to assess the degree to which our objectives have been attained. Unless we know what it is that we expect students to be able to do, we cannot tell whether they can do it or not, whether the "it" is answering questions on a test or doing something for a patient. First, then, we must define our objectives, and this defining should be done at the time we are planning the curriculum.

Preparing Objectives

The school has its program objectives. These are usually statements of goals that describe the product the school intends to produce. They give direction to curriculum planning and should serve as a unifying thread to weave the whole curriculum together. Then, each instructor has certain objectives that direct her teaching. The objectives of the total program must, of course, be reflected in the objectives of each course, but each course is likely to have certain unique objectives. For these reasons, it really isn't possible for an instructor to establish her own objectives apart from those of her colleagues and those of the total program. Together, the faculty is helping to prepare nurses. Unless there is collaboration on objectives, the outcome will be uncertain. Have you ever been to a potluck supper where each person brings whatever she wants to bring? Sometimes you find that six people have brought casseroles, and no one has brought salad. Or everyone has brought dessert, and no one the main course. This is

Mildred E. Katzell, Ph.D., is Director, Division of Measurement and Evaluation Services, National League for Nursing, New York, New York.

why, I guess, at our church, when a potluck dinner is planned, the initial of your last name determines what you will contribute: if it is A to I, you bring a casserole; J to R, a salad; S to Z, dessert.

The collaboration among faculty in curriculum planning is especially important today, when schools of nursing are integrating content throughout the curriculum--subjects like nutrition, pharmacology, growth and development, rehabilitation, teaching, and so on. If there is no collaboration, it is very likely that everyone will assume that someone else is covering certain content, but no one may do it--like no one bringing salad to the potluck supper. Or, alternatively, several people feel comfortable teaching certain content, so it is taught time and time again--like everyone bringing casseroles.

I recall my own freshman year in college, when I was taking Psychology 1, Sociology 1, and Zoology at the same time. Within the space of a single month, I got the Mendelian laws of inheritance in all three courses. Now, I agree that it is important for an educated person to be aware of those laws, but that aware I don't think we needed to be!

Characteristics of Objectives

I don't need to tell you that each nursing education program has many different objectives, and even each course within the curriculum. These may be very specific: The student "administers intramuscular injections"; or very general: "makes nursing judgments." They may be tangible: "specifies the chemical symbol for water"; or intangible: "has positive attitudes toward terminally ill patients." They may be immediate: "states the side effects of Demerol"; or ultimate--so that they will be demonstrated only in the future behavior of the student, when she has become a nurse--such as "checks on the side effects of new or unfamiliar drugs" or "studies procedures before performing new treatments." We may have implicit objectives as well as explicit objectives. All too frequently, it seems, the explicit objectives (those we verbalize or put down on paper) differ markedly from those that are implied by our course content and the things we emphasize in our evaluation of students.

It might be helpful to set down a few of the important characteristics we look for in an objective when it is to be used for purposes of evaluation, whether in the classroom or the clinical laboratory:

1. Each objective should be realistic, appropriate to the program, and theoretically attainable in terms of the school's resources.
2. It should be stated clearly and succinctly, without ambiguity.
3. It should be sufficiently specific to point up both the "what" and the "how" of evaluation--what is to be evaluated, and how the evaluation is to be done.
4. It should represent a single important goal or behavioral outcome.
5. It should be stated in such a way that the desired behavioral change is apparent.
6. It should be stated in such a way that the method of evaluating its attainment is apparent.
7. It should represent observable or measurable behavior.

Next, let us look at how we use our objectives in the evaluation process. To do this,

we need some illustrations of objectives: The student

knows specific facts about drugs.
prepares and administers drugs skillfully.
records pertinent data accurately.
applies scientific knowledge in meeting the physical needs of patients.

Let us take "prepares and administers drugs skillfully" and look at it in terms of the characteristics of objectives. It is realistic; it is appropriate to a diploma program; it is theoretically attainable in terms of the school's resources; it is stated clearly and succinctly, without ambiguity; it is specific enough to indicate what is to be evaluated; it represents a single important goal or behavioral outcome; it is stated in such a way that the desired behavioral change is apparent and it represents observable or measurable behavior. Is it, however, stated in such a way that the method of evaluating its attainment is apparent? We shall see.

How are we going to evaluate the attainment of this objective, for instance: "prepares and administers drugs skillfully"? We need to set down some specific learning outcomes. For example, we could say:

1. Measures dose of insulin accurately, using an insulin syringe.

2. Selects a site for the injection of insulin that has not been used recently.

Now, are you in a better position to say what evaluation techniques can be used? Both of those learning outcomes can be evaluated by means of direct observation of the performance--possibly by the use of pre- and post-conferences. How about students' self-evaluation? The outcome having to do with measuring the dose accurately could also be measured by the use of an objective test question involving the use of a diagram.

Consider some of the other objectives and examples of learning outcomes that guide us toward the evaluation techniques we can use for their assessment. One was "knows specific facts about drugs." One learning outcome would be: "recalls the symptoms of insulin reaction." This can be evaluated by means of an objective test--either multiple-choice or matching questions would do nicely. Another objective was "records pertinent data accurately." Some specific learning outcomes would be:

1. Enters the findings obtained from urine tests for sugar and acetone on the patient's chart immediately upon completion of the tests.

2. Includes the time the urine tests were done.

3. Specifies the exact findings obtained from the urine tests, using symbols or terms appropriate to the method of testing.

Each of these outcomes can be evaluated by direct observation. The second one, "Includes the time the urine tests were done," could also, with a little ingenuity, be tested in an objective test question.

The fourth general objective was "applies scientific knowledge in meeting the physical needs of patients." A specific learning might be: "gives orange juice to a diabetic patient on insulin whom she observes to be semiconscious and mentally confused." This, too, could be evaluated by direct observation, or the knowledge involved could be tested by an objective test question. I have mentioned only a few of the many evaluation tools that can be used. Others included in the objective tests, besides matching and multiple choice, are true-false, completion, combination response, and analogies. Then, there are essay tests, both the restricted-response and the extended-response variety. And

there are a variety of tools for clinical-performance evaluation: anecdotal records, behavior checklists, rating scales, ranking, critical incidents, work samples, self-evaluations, sociograms, and more.

Purposes of Evaluation

Whatever the objective, we must have a way of assessing the extent to which students have attained it. To most instructors, this assessment of attainment serves only the purpose of arriving at a grade for the student. The very least of the many purposes served by the information gained in the evaluation process is that of the pass-fail decision. Probably among the most important purposes of evaluation is that, if carefully planned, it contributes information that is useful for educational as well as administrative actions. For example, evaluation can give answers to these questions:

1. Are the students ready for the next planned learning experiences?
2. Is there some optimal way in which the students can be combined to facilitate the teaching and learning processes?
3. Are there particular aspects of the content that are giving difficulty? To all students? To only some students?
4. Are the students learning what I think I am teaching?
5. Do the students know where their strengths and weaknesses lie?

In addition, the prospect of evaluation stimulates students to be prepared, whether the preparation involves learning content or performing certain procedures in the clinical setting. Evaluation also facilitates learning if it includes evaluation of understanding, application, and interpretation, which require the use and transfer of learning.

In each of these uses of evaluation, you will see that specified objectives are involved, as well as the extent to which they are attained. We can judge whether students are ready for the next planned learning experience only when we know what competencies must precede that experience. If the objective is to combine groups of students into a single group, we will need to know whether they are sufficiently homogeneous in terms of the required competencies to be able to profit from the experience to be provided.

The same comments apply to performance evaluation. Performance evaluation may serve the purpose of assisting in pass-fail and promotion decisions, but, more significantly, it provides an opportunity to find out if particular aspects of the performance are giving difficulty to students, if difficulties are being encountered by all students or only by some, if students are aware of difficulties when they exist, if the problems are in areas that you thought you had taught. Again, the awareness that evaluation is to take place stimulates students to be prepared and to do the best they know how.

The principles of evaluation are equally applicable to performance evaluation. Whatever purposes are to be served by the evaluation of nursing performance, they will be effectively served only if the process is systematic, if the objectives are specified, and if the evaluator assesses the extent to which the objectives have been attained. After all, we cannot judge whether students are ready to function as graduate nurses if we don't know what skills, competencies, attitudes, and abilities we expect the graduate nurse to have.

In addition to the principles that we have already identified for all evaluation, there are certain principles that apply particularly to performance evaluation.

1. We must be knowledgeable about what is to be evaluated. The more you know about the behavior that is to be evaluated, the more valid and penetrating your evaluation is likely to be. The experienced nurse evaluating performance can make more valid and penetrating judgments about performance than can the student or novice.

2. We must evaluate some observable characteristic or measurable aspect. As with essay-test grading, if we set down what it is that we are looking for, we are better equipped to assess the extent to which it has been attained.

3. Performance evaluation involves comparison of observed behavior with previously established standards. The more experience you have in observing certain behaviors, the more clearly will your standards be defined and the better will you be able to distinguish differences in the quality of the performance.

4. Performance evaluation involves interpretations and judgments.

Your ability to evaluate performance improves as you increase your knowledge of what is to be evaluated; as you identify the observable and measurable abilities, behaviors, and skills that are essential to the effective performance of the functions you set out to evaluate; and as you establish standards to be used in defining levels of performance. In performance evaluation, perhaps even more than in classroom evaluation, our objectives must be specified, both for our own use and so that the students are familiar with them. This is desirable for both teaching and evaluation, and it also facilitates the use of the results of the evaluation as a means to improve future performance.

As you well know, evaluating clinical performance is not an easy matter, but it is an essential part of nursing education. A few suggestions may make it a little less difficult.

1. With your colleagues, develop a list of behaviors that you agree are important to student development and direct your observations toward them.

2. With your entire faculty, determine the standards you will use. There are no rules for this. Are you going to compare each student's performance with that of others in the same class, or with her own performance at some earlier point in the program, or with some agreed-upon expectation of all students at a particular stage of the program, or with some other standard?

3. Try to provide comparable experiences for all students, so that their performance is measured by similar "tests."

4. As a faculty, decide how you will handle those questions that inevitably arise when clinical performance is evaluated:

    a. Should we give separate grades in clinical practice and theory?

    b. If we give separate grades, how do we combine them? (How do you combine grades in Chemistry and Psychiatric Nursing?)

    c. Suppose a student has done something wrong and you tell her about it, what do you do if she never has a chance to show what she would do again? Do you grade her on what she did the first time or on what you think she would do the next time?

d. Should we expect different performance by students on different rotation patterns? If a student has completed medical-surgical nursing experience, should our expectations of her performance in obstetric nursing be different from what they would have been had obstetric nursing been her first clinical experience?

e. Should we have different expectations of male and female students? of older students? of others?

5. In rating performance, remember that you can't rate what you can't observe.

6. Rate only those characteristics that are significant in terms of your behavioral objectives (punctuality, neatness, courtesy, other).

7. Don't try to assign grades for clinical performance by making fine discriminations between students. Performance is composed of many factors, and a numerical grade for performance, such as 98 or 82, assumes an ability to discern differences that most instructors lack. It is safe to say that most instructors could differentiate between acceptable and unacceptable performance, and within the acceptable range they could probably identify outstanding performance. But the narrower the classifications of performance for grading purposes, the more likely are we to call for a precision of judgment that is beyond Solomon.

Sources of Error

No matter how carefully we define our objectives and plan our curriculum, many factors can influence the quality of our evaluation.

One of these is the sampling error. If you had asked a different set of questions on your tests, the results very probably would have been different. Or in the clinical setting, if you had observed the student at a different time, you might have observed more or less effective behavior. No matter what method of evaluation you use, you run into sampling errors. Evaluation is based on estimates, not on precise measurements.

You have probably heard of the "halo error." This occurs when the rater tends to assign the same level of rating to all elements of a person's behavior on the basis of a generalized response to that person. Some students have a good deal of verbal fluency and give the impression that they know more than they really do. Some students just rub you the wrong way, and you respond negatively to everything they do. It is hard to recognize good clinical performance on the part of a student who is always late and untidy and has an illegible handwriting. If she is poor in those areas, we tend to rate her poor in everything else: intelligence, knowledge and judgment, observation skills, leadership, dexterity, and all the rest. Or, if she is attractive, neat, punctual, friendly, and docile, then she is bright, capable, a leader, dextrous.

Two other errors made by the rater result from tendencies toward leniency or stringency. The lenient person rates everyone above average; the stringent person rates everyone low. Then there are errors of central tendency: the rater considers everyone as average, avoiding the extremes. To this type of rater, no one is perfect, and no one is all bad.

Errors also occur in evaluation because instructors and students have different expectations, and different instructors have different expectations. If instructors have worked

together to plan curriculum and specify objectives and then have shared the objectives with the students, this kind of problem with evaluation is less likely to arise.

Another major problem is that arising from improper use of the results of evaluation. We have noted that our assessments of students are subject to all kinds of errors, whether the assessment is made on the basis of a term paper or a test or takes place in the clinical setting. Yet, we turn around and attach a degree of precision to our grades that is totally unwarranted. We require one student to withdraw from the school or repeat a course because she got a 68 or a D, while the one who got a 70 or a C remains. Can we really have that kind of confidence in our grading?

Grading

This brings us to grading. There are so many points of view on the subject of grading that we could devote a whole conference to considering them and still not come to any conclusion. But let me identify some of the approaches and raise some of the issues with you so that you can judge which philosophy is closest to your own.

The continuum runs all the way from the 100-point scale down to the pass-fail 2-point scale. According to an article in The New York Times last spring, "there is evidence that perhaps as many as three-quarters of the nation's 3,000 institutions of higher learning have turned to the simple pass-or-fail method of grading as a flexible, though limited, alternative to the traditional A through F marks." Very few schools seem to have adopted the pass-fail system completely; most of those that are using it at all are using it only in a few courses. The justification has been that it encourages students to experiment with broadening courses outside their major field without fear of hurting their academic average and removes the competition for grades that sometimes distorts the learning process and makes the letter grade the goal rather than the learning itself. Also, many instructors find it easier to distinguish between passable and failing performance than to make the finer discriminations required when more categories are used.

Against the pass-fail system, these arguments are advanced: It destroys academic standards, and without the reward of an A or a 98 to strive for, students do the minimum necessary to pass. Resistance is also being encountered from graduate programs, some of which refuse to admit a student with all pass-fail grades, no matter how good the recommendations from faculty members.

We are all familiar with the system of assigning grades of A through F. We also know that these grades mean different things in different contexts to different people. Think of the response to the statement, "She's an A student." Similarly, an F grade has a pretty consistent meaning wherever you find it. But the grades of B, C, and D take on different meanings in different contexts. For a student who has always been an A student, a B is a poor grade; for a C student, a B is an improvement. Some students may be content with a C average, but in graduate school, a C average may not be acceptable. In a particularly difficult course, a student may not be terribly upset if she gets a D when she was afraid she would get an F. Obviously, the meaning of grades is relative--to society, to students, and even to teachers.

Students and teachers sometimes think that grading on the curve means assigning letter grades, and vice versa. In fact, grading on the curve refers to the assignment of letter grades in certain proportions. These proportions are determined by the normal distribution curve, a certain proportion of the group receiving A's; a larger proportion, B's; the

majority C's; fewer, D's; and still fewer, F's. While grading on the curve is statistically reasonable, I wonder if any class of nursing students contains a normal distribution of the abilities associated with achievement in nursing. The students have graduated from high school, usually in the upper half of their class, and most of them took an academic or college preparatory course in high school. Screened for admission, they continue to be screened after admission as they move to successive stages of the program. Consequently, the class is likely to be fairly homogeneous to begin with and to become more so with each passing semester.

This kind of homogeneity creates problems for the instructor who must use a scale of grades. As Robert Ebel has pointed out, it is absurd to give an A for the best achievement in a low-ability group if identical performance would have received a C in a group of higher ability. But it is equally absurd to deny the possibility of recognizing outstanding achievement among low-ability students with a mark of A, or of recognizing, with marks of C, D, or even F, that students of high ability can do what is for them mediocre, poor, or even failing work.

Again, it must be for each faculty as a group to determine the grading practice it will follow. Regardless of the decision, however, if an A is to represent outstanding performance, then its use must be limited to those whose performance is truly outstanding, just as F's are awarded only when performance is truly failing to meet acceptable standards. If all members of a class are fairly comparable in achievement or performance, the faculty should determine in advance whether all should receive average grades or whether, if the entire group is somewhat exceptional, all should receive A's. And suppose they all do receive A's, is an A grade, then, truly a mark of exceptional performance?

Now we turn to numerical grades. In general, we each tend to assume that we are teaching all the content that students should learn in a course and that we are devising tests that validly sample what they have learned of this content. If our assumptions are tenable, then we can comfortably say that any student who receives a grade of 95 to 100 on our tests has demonstrated mastery of 95 to 100 percent of the content. We can readily combine the various numerical marks we have entered in our grade book and come up with a total or an average. We can even convert numerical grades into letter grades or into pass-fail, or, if we want a three-step grading system, into honors-pass-fail. Why isn't this the best system? There are many who would say that it is. For example, one measurement expert points out that you make a far smaller error when you give a student 80 instead of the 75 he might have deserved than you do when you give a student a B instead of a C or a pass grade when he should have failed. But who is to say he should have failed? And who is to say the grade should have been a C, not a B?

Another concept that relates to grading and promotion and about which we are hearing a lot these days is the concept of mastery. I have in my file a newspaper clipping with the heading "Sue the School if the Child Fails." It is an article about accountability in education. The consultant being quoted in the article--his name is Stewart--notes that every time teachers fail a student who has "guessed wrong" on a test, they are downgrading the student for not understanding. A student who answers a question wrong on a test is telling the teacher, "I don't understand." Stewart asks, "If a student came to you after class and said, 'I don't understand,' would you mark a low grade down in your book?" Teachers make students guess what it is they are to learn, he charges, and then fail them for not knowing what it is they are supposed to know. Students should be told what they are supposed to know, and teachers should teach them to learn. Every time you allow a student to progress to the next level with a C grade, you are saying that she knows less than she

really should know to be successful at this higher level. What Stewart recommends is that teachers spell out the objectives of every class and not release any student until she has attained 100 percent of those objectives. Students would receive individualized programs when needed and would learn at their own pace. The implications for grading under this system are probably most similar to those discussed under the pass-fail system except that it is conceivable that students would be encouraged to work on a unit until it was mastered and therefore would not be allowed to fail.

Conclusion

So, where are we now? We have touched on a wide variety of issues related to evaluation and objectives. If you remember nothing else of what I have said today, I would like you to remember this. There are no rules that make evaluation any easier. The one rule that I urge you to follow is that of sharing, cooperating, working together with your colleagues, whether you are formulating objectives, devising curriculum, constructing tests, evaluating performance, or establishing a grading system. You must have the same goals; you must each know what the others are teaching (and not teaching); you should review each other's tests to help improve their quality; you must identify the behaviors that represent satisfactory and unsatisfactory performance. Evaluation isn't easy; it never has been. But like a lot of other things, it is a lot harder when you try to go it alone.

Bibliography

Ebel, Robert L. "Ability vs. Knowledge in Testing Educational Achievement." The National Board Examiner. Philadelphia, National Board of Medical Examiners, May 1969.

_____ Measuring Educational Achievement. Englewood Cliffs, N.J., Prentice-Hall, 1965.

Gronlund, Norman E. Measurement and Evaluation in Teaching, 2d ed. New York, Macmillan, 1971.

Malcolm, Andrew H. "Pass-Fail College Grades Gaining Favor." New York Times, May 1971 (exact date not given).

Moore, Maria J. "Sue School if Child Fails, Consultant Says." Virginian-Pilot, Norfolk, Va., Aug. 6, 1971.

"Sociologist Says Professors Need Help to Evaluate Grades." New York Times, Feb. 7, 1971.

Tyler, Ralph W. "National Assessment: Some Valuable By-Products for Schools." The National Elementary Principal, Vol. XLVIII, No. 6, May 1969.

APPENDIX A

MASTER AGENDA

COUNCIL OF DIPLOMA PROGRAMS
1971 WORKSHOPS

THEME: EVALUATION--AN OBJECTIVE APPROACH

FIRST DAY

Getting Acquainted-- With Each Other With Objectives of Workshop	Chairman and Workshop Participants
Evaluation--An Objective Approach	Speaker
Questions and Answers	Speaker and Participants
Course Objectives and Evaluation Tools An Explanation of the Process An Analysis of Course Objectives and Clinical Laboratory Evaluation Tools	Chairman Work Session

SECOND DAY

Results of Our Analysis	Group Reports
Some Objective Approaches Case Presentations	Panel
Keep Alive--The Objective Approach	Chairman

APPENDIX B

Council of Diploma Programs

National League for Nursing

1971 Workshop: Evaluation--An Objective Approach

WORK SHEET FOR COURSE OBJECTIVES

Course: _____

Number of Objectives: _____

1	2	3
Identify the objectives that specifically differentiate this course from all other courses.	Identify any objectives that are too general to serve as course objectives.	Identify any objectives that are too specific to serve as course objectives.

4	5	6
Identify the objectives that indicate clearly the measurable behaviors expected of the student.	Identify the objectives that do not indicate the measurable behaviors expected of the student. If possible, revise statement so that it does indicate the expected behavior.	In view of your analysis recorded in the other five columns, what recommendations for change in objectives of the course need to be considered by the instructors and/or curriculum committee?

APPENDIX C

Council of Diploma Programs

National League for Nursing

1971 Workshop: Evaluation--An Objective Approach

WORK SHEET FOR CLINICAL LABORATORY EVALUATION TOOL

Course: _____

UNCOMPLETED TOOL

Indicate those statements that specifically relate to the objectives of the course.	Indicate any statement that cannot be directly related to a course objective. If possible, revise statement so that it does relate to an objective.	Identify characteristics that require the evaluator to indicate the degree of attainment of each objective by the student.

COMPLETED TOOL

Identify any statements or items that clearly indicate progress of the student in the course.	Identify any statements or items that show limited relationship to the objectives and student performance in the clinical laboratory.	Analyze the instructions for use of the form and the completed form, and formulate suggestions for improving any aspect of the instructions and/or the tool.